Journal of Econometrics 238 (2024) 105613

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Tail behavior of ACD models and consequences for likelihood-based estimation

Giuseppe Cavaliere [a,b,*], Thomas Mikosch [c], Anders Rahbek [d], Frederik Vilandt [d]

[a] Department of Economics, University of Bologna, Italy
[b] Department of Economics, University of Exeter, UK
[c] Department of Mathematical Sciences, University of Copenhagen, Denmark
[d] Department of Economics, University of Copenhagen, Denmark

ARTICLE INFO

Keywords:
Financial durations
Autoregressive conditional duration (ACD)
Tail index
Quasi maximum likelihood
Mixed normality

ABSTRACT

We establish new results for estimation and inference in financial durations models, where events are observed over a given time span, such as a trading day, or a week. For the classical autoregressive conditional duration (ACD) models by Engle and Russell (1998), we show that the large sample behavior of likelihood estimators is highly sensitive to the tail behavior of the financial durations. In particular, even under stationarity, asymptotic normality breaks down for tail indices smaller than one or, equivalently, when the clustering behavior of the observed events is such that the unconditional distribution of the durations has no finite mean. Instead, we find that estimators are mixed Gaussian and have non-standard rates of convergence. The results are based on exploiting the crucial fact that for duration data the number of observations within any given time span is random. Our results apply to general econometric models where the number of observed events is random.

1. Introduction

In the seminal papers by Engle and Russell (1998) and Engle (2000), autoregressive conditional duration (ACD) models were introduced for modeling durations, or waiting times, between financial events, and to analyze liquidity in financial markets. Financial events are observed over a given period of time, such as a (trading) day, a week, or a year; hence, both the size and the number of durations are random variables. As we demonstrate, the randomness of the number of events has a major impact on asymptotics and inference in dynamic duration models. Moreover, as detailed below, existing results cover alone the case where the number of events is non-random and therefore are not applicable to estimation of ACD models over a given time span. In this paper, we provide the missing asymptotic analysis for likelihood-based estimators. We specifically show that the randomness of the number of events plays a crucial role and leads to a new distributional theory (at non-standard rates of convergence) for likelihood estimators and related test statistics. The derivation of these novel results requires non-standard asymptotic arguments, combining new results on the tail behavior of the durations with renewal theory.

The ACD models are by now quite popular in financial econometrics; see e.g. Hautsch (2012) and Fernandes et al. (2016) for applications and theory in the context of high-frequency data and Pacurar (2008) for a general survey. Applications of dynamic duration models such as the ACD are extensively used also in different areas of economics; see e.g. Hamilton and Jordà (2002) or Aquilina et al. (2022).

* Corresponding author at: Department of Economics, University of Bologna, Italy.
E-mail address: giuseppe.cavaliere@unibo.it (G. Cavaliere).

https://doi.org/10.1016/j.jeconom.2023.105613
Received 13 December 2022; Received in revised form 28 September 2023; Accepted 12 October 2023
Available online 1 December 2023

Let $[0, T]$ denote the observation period, where we observe n event times $\{t_i\}_{i=1}^n$, $0 < t_1 < t_2 < \cdots < t_n \leq T$, with corresponding durations $x_i = t_i - t_{i-1}$, $i = 1, \ldots, n$, $t_0 = 0$. As noted in Engle and Russell (1998), n is the realization at time $t = T$ of the stochastic counting process N_t, $t \geq 0$, given by

$$N_t = \#\{k \geq 1 : t_k = x_1 + \cdots + x_k \leq t\}. \tag{1.1}$$

In particular, the number of events N_T, in the observation period $[0, T]$, is as mentioned a *random variable*.

The most known dynamic duration model is the ACD of Engle and Russell (1998) which in its simplest version (ACD of order one) is given by

$$x_i = \psi_i(\theta)\varepsilon_i, \quad \psi_i(\theta) = \omega + \alpha x_{i-1}, \quad i = 1, \ldots, N_T, \tag{1.2}$$

where $\theta = (\omega, \alpha)'$ and $\psi_i(\theta)$ is the conditional (duration) rate of the ith waiting time x_i, i.e., conditional on $\mathcal{F}_{i-1} = \sigma(x_{i-1}, x_{i-2}, \ldots)$. The innovations $\{\varepsilon_i\}$ are assumed i.i.d., strictly positive, with unit mean, $\mathbb{E}[\varepsilon_i] = 1$. If ε_i is exponentially distributed this is referred to as *exponential* ACD (EACD).

With parameters $\theta = (\omega, \alpha)'$, for $\omega > 0$, $\alpha \geq 0$, and observation period $[0, T]$, the EACD log-likelihood function is given by

$$L_T(\theta) = -\sum_{i=1}^{N_T}\left[\log\psi_i(\theta) + \frac{x_i}{\psi_i(\theta)}\right], \qquad T \geq 0. \tag{1.3}$$

Then $\hat{\theta}_T = \arg\max_\theta L_T(\theta)$ denotes the maximum likelihood estimator (MLE) of θ in the case of i.i.d. exponentially distributed $\{\varepsilon_i\}$, otherwise we refer to it as a quasi maximum likelihood estimator (QMLE).

Engle and Russell (1998) note that the log-likelihood function in (1.3) has the same form as the log-likelihood function for the autoregressive conditional heteroskedastic (ARCH) model with Gaussian innovations, and quote standard asymptotic theory from ARCH models in Lee and Hansen (1994); see also Fernandes and Grammig (2006), Hautsch (2012), Theorem 5.2), Allen et al. (2008) and Sin (2014) for a similar approach to inference. This approach treats N_T as *deterministic*; that is, sampling is by number of durations and not over a fixed, predetermined observation period $[0, T]$. Importantly, the results for deterministic N_T cannot be applied to the case of random N_T, as analyzed here.

To give an idea of the difference in arguments between the two different sampling schemes, a key insight is that the fact that the number of observations N_T is random implies that classical laws of large numbers (LLNs) and central limit theorems (CLTs) are no longer directly applicable to likelihood-related quantities such as score and information. For instance, it is known from renewal process theory (see e.g. (Gut, 2009)) that $N_T \to \infty$ is not sufficient for the LLN or the CLT to apply to series of the form $Y_T = \sum_{i=1}^{N_T}\xi_i$, where both N_T and the random variables $\{\xi_i\}$ are defined in terms of the durations $\{x_i\}$; such series appear repeatedly in the asymptotic theory for ACD. In contrast to the deterministic N_T case, the large sample behavior of Y_T is intimately related to the large sample properties of the counts N_T, which, again, depends on the tail properties (and existence of moments) of the marginal distribution of the stationary and ergodic duration x_i. Such dependence leads to a novel asymptotic theory, based on non-standard arguments.

Specifically, as we show in this paper, the asymptotic theory for the MLE crucially depends on the tail behavior and existence of (unconditional) moments for the ergodic and stationary durations $\{x_i\}$, with the tail behavior characterized by the tail index $\kappa > 0$ of the marginal distribution of x_i; $P(x_i > z) \sim c_\kappa z^{-\kappa}$ as $z \to \infty$ for some constant $c_\kappa > 0$. We show that, while asymptotic normality holds for $\kappa > 1$, or equivalently, when the durations have finite mean, asymptotic normality breaks down for $\kappa < 1$. Notably, in finite sample, the Gaussian asymptotic approximation is poor for the case of infinite variance $\kappa < 2$ and indeed invalid for the case of infinite mean where $\kappa < 1$. This is a crucial fact, given that a wide range of tail indices is witnessed in empirical applications on duration data. Thus, for example, Hill estimation of κ yields $\hat{\kappa} = 2.1 > 2$ for the (diurnally-adjusted) IBM transaction data analyzed in Engle and Russell (1998) and $\hat{\kappa} = 2.5 > 2$ on the durations between tweets in Cavaliere et al. (2023). Moreover, $\hat{\kappa} = 1.4 \in (1, 2)$ for the DJIA data from Embrechts et al. (2011), while $\hat{\kappa} = 0.81 < 1$ on diurnally-adjusted durations of GameStop stock intraday trades (January 1 to February 24, 2021); see Fig. 1 for plots of durations and the corresponding histograms.[1]

A preview of our results is as follows. In classic settings, with $\{x_i\}$ i.i.d. with finite mean, the number of events per unit of time N_T/T converges almost surely to a strictly positive constant, in which case LLNs and CLTs for $\sum_{i=1}^{N_T}\xi_i$ can usually be verified; see e.g. Gut (2009) for a survey. In the ACD setting, whether this holds depends on the tail index κ. On the one hand, if $\kappa > 1$, hence $\mathbb{E}[x_i] = \mu \in (0, \infty)$, and N_T/T is such that $N_T/T = 1/\mu + o(1)$, a.s. However, even in this simpler case, existing (renewal) theory does not include stationary and ergodic x_i, and we provide the needed extensions to the theory here. On the other hand, if $\kappa < 1$, hence $\mathbb{E}[x_i] = \infty$, then N_T/T converges (a.s.) to zero as $T \to \infty$ and, in particular, the CLT does not apply to $\sum_{i=1}^{N_T}\xi_i$. New tools are required for the asymptotic theory and, specifically, we establish the novel result that N_T/T^κ converges in distribution to a random variable with an unfamiliar distribution, and for which we provide an explicit expression in terms of a κ-stable random variable.

These convergence results for N_T are essential for establishing the asymptotic distribution of the QMLE. Specifically, we show that, provided $\mu = \mathbb{E}[x_i] < \infty$, $\hat{\theta}_T - \theta_0$ (with θ_0 denoting the true value) is indeed asymptotically Gaussian when normalized by the standard deterministic $\sqrt{T}$-rate. However, while $\mathbb{E}[x_i] < \infty$ is indeed sufficient for $\sqrt{T}$-convergence to the Gaussian distribution, the quality of the Gaussian approximation in finite samples is demonstrated to be very poor when $\mathbb{E}[x_i^2] = \infty$, or $\kappa < 2$, and deteriorating as the tail index κ approaches one. Hence even when $\mathbb{E}[x_i] < \infty$ these results question the usefulness of the $\sqrt{T}$-Gaussian approximation for likelihood estimators in ACD models. In the case $\kappa < 1$, the fact that N_T/T^κ converges in distribution

[1] See, e.g., Drees et al. (2000) for guidance on how to estimate κ using the Hill-estimator; in particular on selecting the number of order statistics.

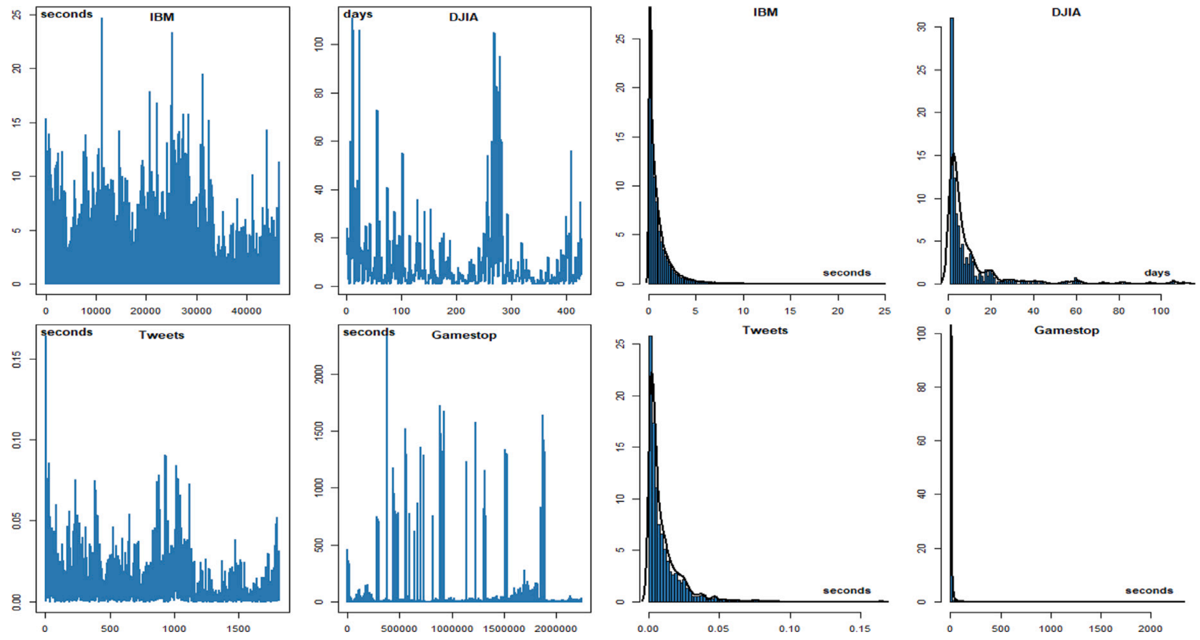

Fig. 1. Left panel: duration time-series $\{x_i(j)\}_{i=1,\ldots,N_T}$, $j = 1,2,3,4$. Right panel: corresponding histograms. Duration data sets: ($j = 1$) IBM transaction data analyzed in Engle and Russell (1998), ($j = 2$) durations between tweets in Cavaliere et al. (2023), ($j = 3$) DJIA data from Embrechts et al. (2011), and ($j = 4$) durations between GameStop stock trades from January to February 2021 (LOBSTER database).

– and not in probability – to a non-standard random variable implies that the derivation of the limiting distribution of $\hat{\theta}_T - \theta_0$ is non-standard. In particular we show that the information is random in the limit, and that this results in a limiting mixed Gaussian distribution of $\hat{\theta}_T - \theta_0$ with a convergence rate which depends on the value of tail index $\kappa < 1$. A further, novel result that follows from our results is that the t-ratio for (univariate) hypotheses on θ is asymptotically normal provided $\kappa > 1$ or $\kappa < 1$. The local power function of the test, however, crucially depends on κ. The case $\kappa = 1$ is not covered by our theorem, and hence particular attention should be paid to applications where estimated parameters are close to the boundary case $\mathbb{E}[x_i] = \infty$.

To sum up, our results show that, in contrast to ARCH models where the marginal distribution of the data does not play any role in the asymptotic theory, for ACD models this is indeed crucial, as the tail index of the duration determines the speed of convergence of the estimators as well as their asymptotic distribution. As already mentioned this is of empirical relevance, as both the case of infinite and finite mean durations ($\kappa < 1$ and $\kappa > 1$, respectively) are found in applications. Moreover, our findings are not specific to ACD models, but apply to general econometric method where the number of observations over a given time span needs being treated as a random process; see also Section 4.

The paper is structured as follows. In Section 2 we discuss the tail behavior of ACD processes, and provide new results for the related counting process N_T, $T \geq 0$. In Section 3 we present the main asymptotic theory. A discussion about the implications for inference and some concluding remarks are given in Section 4. All proofs are provided in the Appendix. In the following, '$\overset{p}{\to}$', '$\overset{\text{a.s.}}{\to}$', and '$\overset{d}{\to}$' refer to convergence in probability, almost surely and in distribution, respectively, in all cases when $T \to \infty$. A generic element of a strictly stationary sequence $\{y_i\}$ is denoted by y.

2. Preliminaries

In this section we derive the required results on the tail properties of the durations and on the asymptotic behavior of the random number of durations N_T. Results of this kind are neither present nor required in the classical ARCH case, where N_T is deterministic.

2.1. Tail behavior of the ACD

We consider the sequence $x_i = \psi_i \epsilon_i$, $i \in \mathbb{Z}$, given as the solution to the ACD Eq. (1.2), and state explicit conditions for stationarity and geometric ergodicity of $\{x_i\}$ as well as for power-law tails of x with index κ. The range of the values κ will be crucial for our asymptotic theory.

The results are initially stated for general positive i.i.d. distributed innovations $\{\epsilon_i\}$.

Lemma 2.1 (*ACD Properties*)**.** *Consider* $\{x_i\}$ *given by* (1.2) *with a strictly positive i.i.d. sequence* $\{\epsilon_i\}$ *with density* f_ϵ*, and for which* $\mathbb{E}[\epsilon] = 1$ *and* $s^2 = \mathbb{E}[\epsilon^2] < \infty$*. Then* $\{x_i\}$ *is geometrically ergodic and has a stationary representation for* $\alpha \in (0, a_u)$, $a_u = \exp(-\mathbb{E}[\ln(\epsilon)]) > 1$.

Moreover, if the unique positive solution $\kappa = \kappa(\alpha) > 0$ *to the equation* $\mathbb{E}[(\alpha\varepsilon)^{\kappa}] = 1$ *exists, then* $\mathbb{P}(x > z) \sim c_{\kappa} z^{-\kappa}$, $z \to \infty$, *for some positive constant* c_{κ} *given in* (A.2). *In particular, we have*

$$\begin{cases} 2 < \kappa < \infty\,, & \text{for } \alpha \in (0, 1/s) \\ \kappa = 2\,, & \text{for } \alpha = 1/s \\ 1 < \kappa < 2\,, & \text{for } \alpha \in (1/s, 1) \\ \kappa = 1\,, & \text{for } \alpha = 1 \\ 0 < \kappa < 1\,, & \text{for } \alpha \in \left(1, a_u\right) \end{cases}$$

The results in Lemma 2.1 complement existing results on ARCH processes; see e.g. Embrechts et al. (1997), Buraczewski et al. (2016), Carrasco and Chen (2002), Fernandes and Grammig (2006), and allow in particular one to assess the existence of moments of ACD processes. Thus, we find that for $\alpha < 1$ the mean is finite, $\mathbb{E}[x] < \infty$, while the variance $\mathbb{V}[x]$ is finite for α in the smaller region $(0, 1/s)$, where $s^2 = \mathbb{E}[\varepsilon^2]$. For $1 < \alpha < a_u$, while $\{x_i\}$ is a strictly stationary and geometrically ergodic sequence, only fractional moments (of order less than one) of x are finite.

Next, we consider the benchmark model where ε is exponentially distributed (EACD).

Lemma 2.2 (*EACD Properties*)**.** *Consider* $\left\{x_i\right\}$ *given by* (1.2) *with an i.i.d. sequence* $\{\varepsilon_i\}$ *exponentially distributed with* $\mathbb{E}[\varepsilon] = 1$. *The Eq.* (1.2) *has a strictly stationary geometrically ergodic solution* $\{x_i\}$ *if and only if* $\alpha \in [0, a_u)$, *with* $a_u = \exp(\gamma) \simeq 1.8$, *where* γ *is Euler's constant. The remaining results in* Lemma 2.1 *hold with* $s^2 = \mathbb{E}[\varepsilon^2] = 2$; *with* Γ *denoting the Gamma function, the equation* $\mathbb{E}[(\alpha\varepsilon)^{\kappa}] = 1$ *has in this case a unique implicit solution given by*

$$\alpha = [\Gamma(\kappa + 1)]^{-1/\kappa}. \tag{2.1}$$

In particular, we observe the surprisingly simple explicit relationship between α and $\kappa = \kappa(\alpha)$ in (2.1) which comes from the properties of the exponential distribution. Such a simple relationship does not exist for general distributions of ε and more general functional forms of ψ_i.

2.2. Asymptotics for the ACD counting process

In Lemma 2.3 below we collect some novel asymptotic results for the counting process N_T, $T \geq 0$, which are needed for the asymptotic analysis of the QMLE of the ACD process.

Our results are general and of independent interest, in particular as the dependence of the durations sequence is an uncommon condition in the literature on renewal theory; there it is typically assumed that the durations are i.i.d. or at most m-dependent (e.g. finite moving average); see e.g. Gut (2009), Janson (1983). Moreover, and also new with respect to existing theory, we present results for the convergence of the counting process N_T when durations have a tail index $\kappa < 1$.

Recall initially that N_T is defined in terms of the dependent sequence $\left\{x_i\right\}$, cf. (1.1), with x_i defined in (1.2). As in Lemma 2.1 we consider here the general case of positive i.i.d. innovations $\left\{\varepsilon_i\right\}$ with unit mean. The following result provides convergence rates for N_T as $T \to \infty$.

Lemma 2.3. *Consider a strictly stationary geometrically ergodic positive solution* $\{x_i\}$ *to* (1.2) *with tail index* $\kappa > 0$ *and* $\{\varepsilon_i\}$ *an i.i.d. sequence with* $\varepsilon > 0$, $\mathbb{E}[\varepsilon] = 1$. *Then the following results hold for the counting process* N_T, $T \geq 0$, *defined in* (1.1).

(i) *For* $\kappa > 1$,

$$N_T/T \overset{\text{a.s.}}{\to} 1/\mu\,, \quad \text{where } \mu = \mathbb{E}[x] < \infty.$$

(ii) *For* $\kappa > 2$, *the CLT holds:*

$$T^{1/2}(N_T/T - 1/\mu) \overset{d}{\to} N(0, \sigma^2/\mu^3),$$

where $\sigma^2 = \mathbb{E}[(1 + T_\infty)^2 - T_\infty^2]\mathbb{V}[x]$ *and* $T_\infty = \sum_{i=1}^{\infty} \alpha^i (\prod_{j=1}^{i} \varepsilon_j)$. *For* $\kappa = 2$ *the CLT holds with normalization* $c\sqrt{T \ln T}$ *for some positive constant* c, *and a standard normal limit distribution.*

(iii) *For* $1 < \kappa < 2$,

$$T^{(\kappa-1)/\kappa}(N_T/T - 1/\mu) \overset{d}{\to} \gamma_\kappa = \left(c_\kappa \mathbb{E}\left[(1 + T_\infty)^\kappa - T_\infty^\kappa\right]/\mu\right)^{1/\kappa} \eta_\kappa,$$

where η_κ *is a totally skewed to the right* κ*-stable random variable whose characteristic function is given in* (A.6), *and* c_κ *is defined in* (A.2) .

(iv) *For* $0 < \kappa < 1$, $N_T/T \overset{\text{a.s.}}{\to} 0$ *and*

$$\frac{N_T}{T^\kappa} \overset{d}{\to} \lambda_\kappa = \left(c_\kappa\, \mathbb{E}[(1 + T_\infty)^\kappa - T_\infty^\kappa]\right)^{-1} \eta_\kappa^{-\kappa}, \tag{2.2}$$

where η_κ *is a totally skewed to the right* κ*-stable random variable whose characteristic function is given in* (A.6), *and* c_κ *is defined in* (A.2) .

It is worth noticing that for all cases (i)–(iv), $N_T \to \infty$ a.s. as a consequence of $x > 0$. However, the convergence rates are quite distinct, depending on κ. Thus, $N_T/T \to 1/\mu$ a.s. for $\kappa > 1$, while, for $\kappa < 1$, N_T/T^κ converges in distribution to the positive random variable λ_κ, and in particular, $N_T/T \to 0$ a.s. For $\kappa > 2$, $N_T/T - 1/\mu$ satisfies the CLT with standard $\sqrt{T}$-rate, while for $1 < \kappa < 2$, the rate $T^{(\kappa-1)/\kappa}$ gets slower as κ gets closer to 1. We also note that the κ-stable limiting random variable η_κ has power-law tail with index κ. Importantly, for the novel result on the distributional convergence of N_T/T^κ for $\kappa < 1$, the limiting variable λ_κ has exponentially decaying tails; cf. Theorem 2.5.2 in Zolotarev (1986).

3. Asymptotic theory for the QMLE

In this section we derive the asymptotic properties of the (Q)MLE $\hat{\theta}_T = \arg\max_\theta L_T(\theta)$, with $L_T(\theta)$ defined in (1.3). Note that, as is common practice, $L_T(\theta)$ is defined without the additional term corresponding to the fact that no events are observed in the end-period $(t_{N_T}, T]$. We show in Appendix B that this term has no influence on the asymptotic results.

We start in Section 3.1 by discussing the behavior of the score and information, which is key to the asymptotic analysis. Then, in Section 3.2, we present the main results on the asymptotic behavior of $\hat{\theta}_T$.

3.1. Convergence of the score and information

With the likelihood function $L_T(\theta)$ as given in (1.3), the corresponding score and information functions, evaluated at the true value $\theta = \theta_0$, are given by

$$S_T = \left.\frac{\partial L_T(\theta)}{\partial\theta}\right|_{\theta=\theta_0} = \sum_{i=1}^{N_T} \xi_i, \quad \xi_i = (\varepsilon_i - 1)\,\mathbf{v}_i, \qquad \mathbf{v}_i = (1, x_{i-1})'/\psi_i, \tag{3.1}$$

$$I_T = -\left.\frac{\partial^2 L_T(\theta)}{\partial\theta\partial\theta'}\right|_{\theta=\theta_0} = \sum_{i=1}^{N_T} \zeta_i, \qquad \zeta_i = (2\varepsilon_i - 1)\,\mathbf{v}_i\mathbf{v}_i', \tag{3.2}$$

where $\psi_i = \psi_i(\theta_0)$. In what follows, we always assume that the conditions of Lemma 2.1 are satisfied. In particular, (1.2) has a strictly stationary geometrically ergodic solution $\{x_i\}$ with tail index $\kappa > 0$.

Consider first the case $\kappa > 1$, where we have the following result on the large sample behavior of S_T and I_T at standard rates of convergence.

Lemma 3.1. *Assume that for $\theta_0 = (\omega_0, \alpha_0)'$, $\omega_0 > 0$ and $\alpha_0 > 0$ such that $\{x_i\}$ in (1.2) is stationary and ergodic, with $\kappa > 1$. With $\tau = \mathbb{V}[\varepsilon]$ and $\Omega = \mathbb{E}[\mathbf{v}_1\mathbf{v}_1']$ we have*

$$T^{-1/2}S_T \overset{d}{\to} (\tau\Omega/\mu)^{1/2}\,\mathbf{Z} \text{ and } T^{-1}I_T \overset{\text{a.s.}}{\to} \Omega/\mu, \tag{3.3}$$

where $\mathbf{Z}$ is a bivariate standard Gaussian vector. Moreover, $N_T^{-1}I_T \overset{\text{a.s.}}{\to} \Omega$.

Next turn to the case $\kappa < 1$ such that $\mathbb{E}[x] = \infty$. As shown in the next, the score and information converge at slower rates than usual. More specifically, turning to the information, it follows by Lemma 2.3 that $N_T/T^\kappa \overset{d}{\to} \lambda_\kappa$ and (see the proof of Lemma 3.2 below) $N_T^{-1}I_T \overset{\text{a.s.}}{\to} \Omega$. Hence,

$$T^{-\kappa}I_T \overset{d}{\to} \lambda_\kappa\Omega. \tag{3.4}$$

That is, the rate of convergence is indeed slower than standard when $\kappa < 1$, and the observed information is random in the limit due to the random variable λ_κ. Similarly, non-standard convergence rates as a function of κ also apply to the score as we state the following lemma for the EACD.

Lemma 3.2. *Assume that for $\theta_0 = (\omega_0, \alpha_0)'$, $\omega_0 > 0$ and $\alpha_0 > 0$, such that $\{x_i\}$ in (1.2) is stationary and ergodic with $\kappa < 1$, and ε_i exponentially distributed with $\mathbb{E}[\varepsilon] = 1$. With Ω defined in Lemma 3.1 we have*

$$(T^{-\kappa/2}S_T, T^{-\kappa}I_T) \overset{d}{\to} \left((\lambda_\kappa\Omega)^{1/2}\mathbf{Z}, \lambda_\kappa\Omega\right),$$

where $\mathbf{Z}$ a bivariate standard Gaussian vector, independent of λ_κ defined in (2.2). Moreover, $N_T^{-1}I_T \overset{\text{a.s.}}{\to} \Omega$.

3.2. Limit theorems for the QMLE

We are now in the position to state the asymptotic distribution of the QMLE $\hat{\theta}_T$ of θ. As for the score, the limit behavior of the QMLE depends on the tail behavior of the durations $\{x_i\}$. As mentioned, the influence of the right power-law tail of x on the QMLE is in contrast to QMLE theory for ARCH and GARCH processes where the shape of the unconditional distribution does not matter. We show here that for ACD processes the power-law tails determine the limiting distribution of the QMLE $\hat{\theta}_T$ as well as the rate of convergence. This result appears surprising, given that, apart from the random summation index, the ACD (log-)likelihood function is identical to the ARCH Gaussian likelihood function.

Specifically, while $\sqrt{T}$-asymptotic normality holds when the tail index κ is above one, when $\kappa < 1$, the speed of convergence and the limiting distribution are non-standard. In particular, for the case $\kappa > 1$ the following result holds.

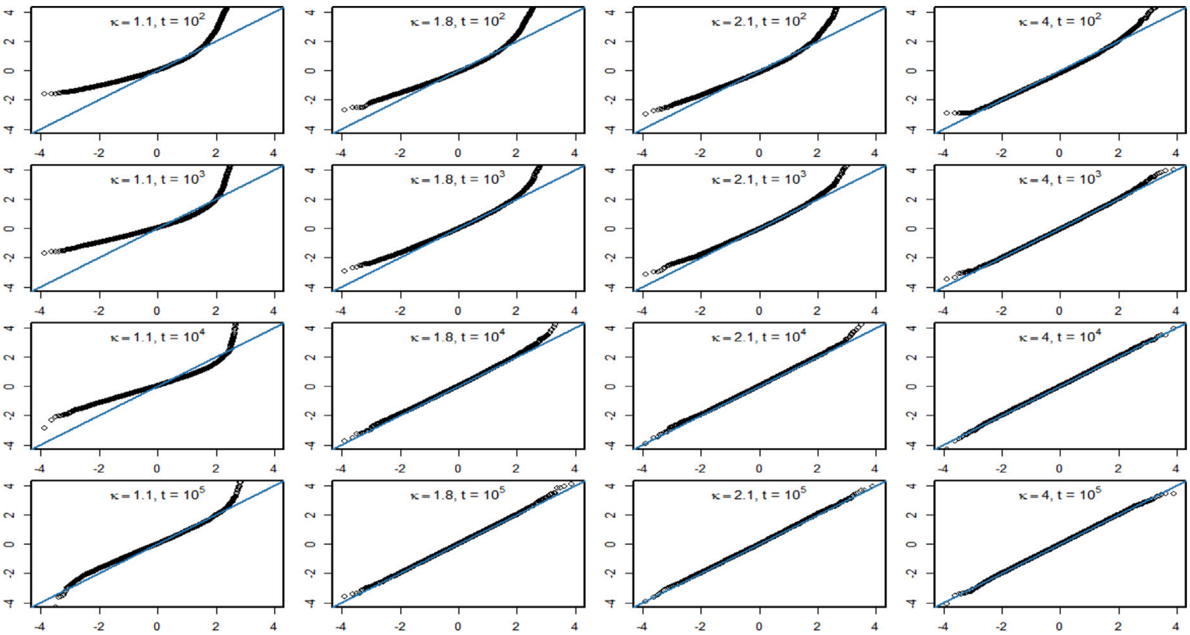

Fig. 2. Finite mean case, $\kappa > 1$. Q-Q plots of $T^{1/2}(\hat{\alpha}_T - \alpha_0)/\sigma_\alpha$, with σ_α^2 the asymptotic variance of $\hat{\alpha}_T$, against the $N(0,1)$ distribution for different values of T (rows) and $\kappa_0 = \kappa(\alpha_0)$ (columns), finite mean case ($\kappa > 1$). $M = 10^4$ Monte Carlo replications.

Theorem 3.1. *Under the assumptions of Lemma 3.1, with probability tending to one, there exists a local maximum $\hat{\theta}_T$ of $L_T(\theta)$ which satisfies $\hat{\theta}_T \xrightarrow{p} \theta_0$ and $\partial L_T(\theta)/\partial\theta|_{\theta=\hat{\theta}_T} = 0$. Moreover,*

$$T^{1/2}(\hat{\theta}_T - \theta_0) \xrightarrow{d} (\Omega/\mu)^{-1/2}\,\tau^{1/2}\mathbf{Z}, \tag{3.5}$$

with $\mathbf{Z}$ a bivariate standard Gaussian vector, $\Omega = \mathbb{E}[\mathbf{v}_1\mathbf{v}_1']$, $\mu = \mathbb{E}[x]$ and $\tau = \mathbb{V}[\varepsilon]$.

Theorem 3.1, which is based on combining classic likelihood expansions with the results for a random summation index N_T in Section 2, shows that asymptotic normality at the $\sqrt{T}$-rate holds even if the durations have infinite variance, $\mathbb{E}[x^2] = \infty$. Note in this respect that the difference between the asymptotic variance in (3.5) and the classic ARCH QMLE asymptotic variance is the term $\mu = \mathbb{E}[x_i] = \omega_0(1-\alpha_0)^{-1}$. To see this, note that (3.5) can be rewritten as

$$T^{1/2}(\hat{\theta}_T - \theta_0) \xrightarrow{d} N(0, \mu V),\; V = \Omega^{-1}\tau$$

where V is identical to the classical ARCH QMLE asymptotic variance; see, e.g., Theorem 3 in Lee and Hansen (1994), with $A_0 = \tau\Omega/\mu$ and $B_0 = \Omega/\mu$ using (3.3).

Importantly, the asymptotic approximation deteriorates as the tail index κ approaches one. This reflects the fact that the asymptotic results for the QMLE when $\kappa > 1$ are essentially derived by replacing the random indices N_T in the likelihood function (and its derivatives) by the deterministic function T/μ; this replacement, however, happens with a much larger error for $\kappa \in (1,2)$ than in the finite variance case ($\kappa > 2$), due to slow convergence rates of $N_T/T - 1/\mu$ (cf. Lemma 2.3) and the widespread limit distribution.

As an explanation to the fact that while the rate of convergence is standard $\sqrt{T}$ for all $\kappa > 1$, the convergence to the Gaussian limit for $\kappa \in (1,2)$ slows down when compared to the (finite variance) case $\kappa > 2$, consider here the score S_T. By Lemma 2.3(iii), for $\kappa \in (1,2)$

$$T^{-1/2}S_T = [T^{(1-\kappa)/\kappa}\hat{\gamma}_\kappa/(2\sqrt{\mu}) + 1/\sqrt{\mu} + o_p(1)]\hat{Z}_T,$$

where $\hat{\gamma}_\kappa = T^{(\kappa-1)/\kappa}\left(N_T/T - 1/\mu\right) \to_d \gamma_\kappa$ (a κ-stable random variable) and $\hat{Z}_T = N_T^{-1/2}S_T \to_d (\tau\Omega)^{1/2}\,Z$. Additionally, γ_κ is non-standard distributed with a power-law tail with index κ, and is more widespread as κ diminishes. Thus, as κ approaches one, $T^{(1-\kappa)/\kappa}\hat{\gamma}_\kappa/(2\sqrt{\mu})$ converges to zero at a slower speed, and the convergence (in distribution) of $T^{-1/2}S_T$ to the Gaussian distribution slows down. This is in contrast to the case $\kappa > 2$, where by Lemma 2.3 (ii), $T^{-1/2}S_T = [T^{-1/2}\hat{\eta}_T/(2\sqrt{\mu}) + 1/\sqrt{\mu} + o_p(1)]\hat{Z}_T$, with $\hat{\eta}_T = T^{1/2}(N_T/T - 1/\mu)$ asymptotically Gaussian. In particular, the rate is independent of κ in this case.

We illustrate this in Fig. 2, where we report Q-Q plots of $T^{1/2}(\hat{\alpha}_T - \alpha_0)/\sigma_\alpha$, with σ_α^2 is the asymptotic variance of $\hat{\alpha}_T$, against the Gaussian distribution when the data follows an EACD process with $\mathbb{E}[x] = 1$, for different values of T and κ. The figure clearly shows how the tail index of the durations influences the quality of the Gaussian approximation in finite time intervals. It can also be seen that as κ gets closer to one, the asymptotic approximation requires larger values of T to be accurate. Unreported simulations show

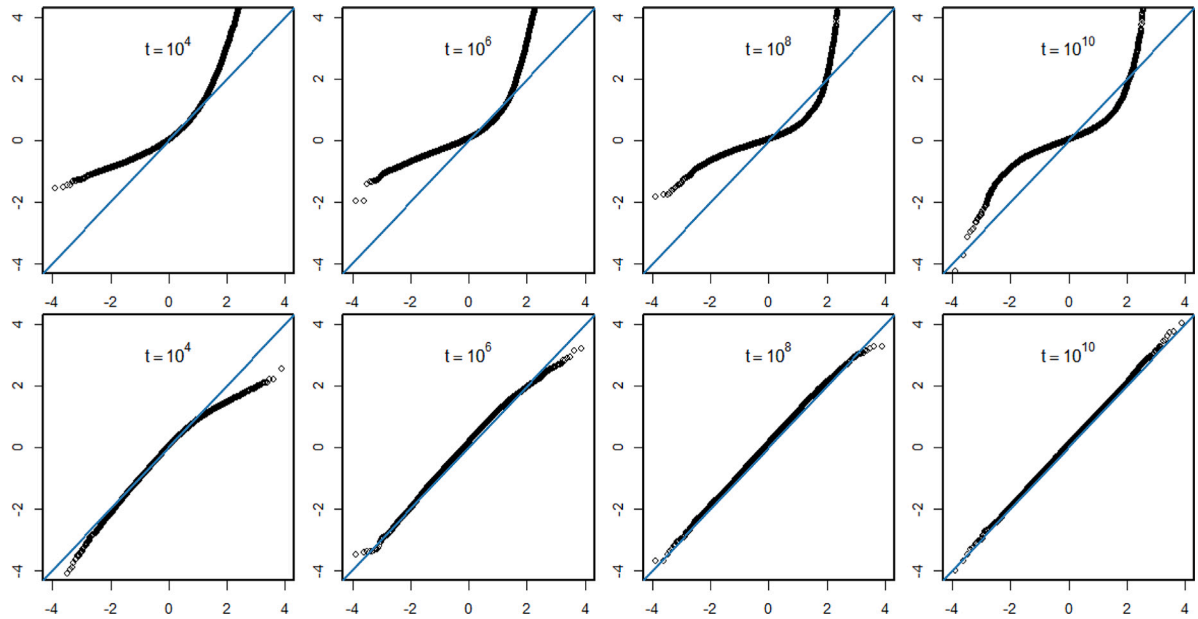

Fig. 3. Infinite mean case, $\kappa = 0.5$, for different values of T with ω_0 selected such that the median of x_i is about one. Q-Q plots against the $N(0,1)$ distribution. Upper panel: $T^{\kappa/2}(\hat{\alpha}_T - \alpha_0)$ (normalized by empirical standard deviation across Monte Carlo replications). Lower panel: t-ratios. $M = 10^4$ Monte Carlo replications.

that if ε follows the Weibull distribution (with different shape parameters), the results in Fig. 2 for the QMLE remain unchanged; the same holds when the Burr distribution (which generates non-monotonic hazard rates, see (Grammig and Maurer, 2000)) is used, albeit larger sample sizes are required.

For $\kappa < 1$, as previously emphasized, $\hat{\theta}_T - \theta_0$ is not asymptotically Gaussian distributed.

Theorem 3.2. *Under the assumptions of Lemma 3.2, with probability tending to one, there exists a local maximum $\hat{\theta}_T$ of $L_T(\theta)$ which satisfies $\hat{\theta}_T \xrightarrow{p} \theta_0$ and $\partial L_T(\theta)/\partial\theta|_{\theta=\hat{\theta}_T} = 0$. Moreover,*

$$T^{\kappa/2}(\hat{\theta}_T - \theta_0) \xrightarrow{d} \left(\lambda_\kappa \Omega\right)^{-1/2} \mathbf{Z}, \tag{3.6}$$

with $\Omega = \mathbb{E}[\mathbf{v}_1\mathbf{v}_1']$ and $\mathbf{Z}$ a bivariate standard Gaussian vector, independent of λ_κ defined in (2.2).

Thus, for $\kappa < 1$ the estimators are asymptotically mixed Gaussian; moreover, the rate of convergence $T^{\kappa/2}$ is lower than the standard $T^{1/2}$ rate and depends on the value of κ. The non-Gaussianity in (3.6) is clearly illustrated in the upper panel of Fig. 3, which reports Q-Q plots of $T^{\kappa/2}(\hat{\alpha}_T - \alpha_0)$ (normalized by its empirical standard deviation). The lower panel reports t-ratios which by Corollary 4.1 below are asymptotically Gaussian.

4. Discussion and implications for inference

In the previous section we have shown that, for the case of a finite mean of the durations, the (Q)MLE is indeed asymptotically normal at the standard $\sqrt{T}$-rate while, for the case of infinite mean, the limiting distribution is a mixture and convergence attains at a lower rate.

In terms of inference, from Theorems 3.1 and 3.2 we can derive the following new result, which shows that t-ratios are asymptotically standard Gaussian distributed, irrespective of the tail index of the durations κ being above or below unity. That is, while κ affects the distributional theory for of (Q)MLE, asymptotic inference based on t-tests (or likelihood ratio tests) is standard, and asymptotic validity holds irrespective of the tail index of the marginal distribution of the durations $\{x_i\}$.

Corollary 4.1. *Under the assumptions of Theorem 3.1 ($\kappa > 1$) or under the assumptions of Theorem 3.2 ($\kappa < 1$), it holds that the t-ratio $t_T = \text{se}_T^{-1}(\hat{\alpha}_T - \alpha_0)$, where $\text{se}_T^2 = \hat{\tau}_T a_T$, a_T being the entry of $I_T(\hat{\theta}_T)^{-1} = (-\frac{\partial^2 L_T(\theta)}{\partial\theta\partial\theta'}\Big|_{\theta=\hat{\theta}_T})^{-1}$ corresponding to α, is standard normal as $T \to \infty$, provided $\hat{\tau}_T$ is a consistent estimator of τ.*

Convergence of the t-ratios to the Normal distribution is illustrated in the lower panel of Fig. 3, where Q-Q plots of t_T against the $N(0,1)$ distribution are reported for increasing sample sizes and for $\kappa < 1$. The figure clearly shows that extremely large observation periods are required for the normal asymptotic approximation to be accurate. This implies that in empirical applications, and differently from inference in ARCH models, inspection of the tails of the marginal distribution of the data is a key step to be taken prior to any empirical analysis.

Finally, we note that, in terms of theory, the result in Corollary 4.1 for $\kappa < 1$ is similar to the mixed Gaussian limit results as employed, e.g., in the cointegration analysis of non-stationary variables; see Johansen (1991) and Phillips (1991).

5. Conclusions

Our results demonstrate the sensitivity of the limiting distribution of the QMLE in ACD models to the tail behavior, or equivalently finiteness of unconditional moments, of the durations. The theory in Engle and Russell (1998) states that the estimators are asymptotically Gaussian under the assumption of stationarity and ergodicity of the durations and, hence, as independent of the duration's tail behavior. Put differently, sampling over a fixed period of time (hence implying a random number of events) leads to a new non-standard theory, while sampling over a fixed number of events (hence implying a random length of observation period) leads to standard theory as in ARCH models.

As is well-known, estimation of the tail index κ is not an easy task, even with i.i.d. data. Our paper suggests that estimation of tail indices for dependent time series such that for ACD models is a key issue. At present, we suggest to test for (in)finite moments in the unconditional distribution of the durations by using results in, e.g., Trapani (2016) and Francq and Zakoïan (2022). Also note that the bootstrap may be an advisable tool in this framework. For the case $\kappa > 1$, results in Cavaliere et al. (2023, Section 6) can be applied such that a fully recursive bootstrap consistently estimate the asymptotic distribution of the MLE of the exponential ACD model. It is worth noticing that these results do not directly extend to cases where $\kappa \leq 1$, nor to QMLE. Unreported, Monte Carlo simulations seem to indicate that the recursive bootstrap is indeed applicable even for $\kappa < 1$.

All results have been stated for the tail index $\kappa < 1$ (infinite mean) or $\kappa > 1$ (finite mean). A tail index $\kappa = 1$, which corresponds to the 'integrated' ACD, is not covered by our theory and in this case the asymptotic distribution of the QMLE remains unknown.

We note that our theory can be extended to more general ACD models, in particular to the much applied ACD(1,1) model where $\psi_i = \omega + \alpha x_{i-1} + \beta\psi_{i-1}$ – that is, the ACD analogue of the GARCH(1,1) – as well as its extensions. We have refrained from doing so here to keep the presentation simple, and thereby focus on the main new insights. Note here that, as for GARCH(1,1) models, the representation of ψ_i as a stochastic recurrence equation as in Buraczewski et al. (2016) can be exploited (see also (Mokkadem, 1990)).

Finally, it is worth noticing that our findings and arguments, are not specific to the models for time series of durations. Indeed, they apply to any econometric method where the number of observations needs being treated as random. For example, asymptotic theory for daily realized volatility, see Li et al. (2014), treats summations such as $\sum_{i=1}^{N_T}(p_{t_i} - p_{t_{i-1}})^2$, where p_t is the (log-)price at time t. Since the number of trades within a day, N_T, is random, our results could be applied to cases where $x_i = t_i - t_{i-1}$ have heavy tails. Note in this context that a key challenge is to establish joint convergence *in distribution* of the score and information with N_T random and $N_T \to \infty$ a.s. For the case of $\kappa > 1$, this can be based on standard asymptotic arguments. For the case of $\kappa < 1$, this is more delicate and non-standard arguments are needed, as is clear from the proofs of Lemma 3.2 and Theorem 3.2. The key difference between the two cases is that in the former case, $N_T/T \to c > 0$ in probability, while in the latter case, $N_T/T^\kappa \to \eta > 0$, in *distribution*, in particular with η random. Li et al. (2014) (see also Bollerslev et al., 2020, 2022a,b) apply stable convergence in distribution (cf. Theorem 1 in Li et al. (2014)). In essence, this would be equivalent in our setting to establish the much stronger condition that $N_T/T^\kappa \to \eta > 0$, in *probability* (as opposed to distribution). For the case of $\kappa > 1$, the two approaches are equivalent.

Acknowledgments

We thank Torben Andersen (co-editor), an associate editor and two anonymous referees for several important comments and suggestions on a previous draft of this work. We are also grateful to Federico Bandi, Matias Cattaneo, Marcelo Fernandes, Nikolaus Hautsch, Marcelo Medeiros and Stefan Voigt for comments and discussions. We also thank participants at the SoFiE 2022 conference (U Cambridge), the 2022 Aarhus Workshop in Econometrics (Aarhus U), the 3rd High Voltage Econometrics meeting, the 2022 Bologna/Rome-Waseda Time Series Workshop, the 10th Italian Congress of Econometrics and Empirical Economics (U Cagliari), the 2023 Financial Econometrics Conference (Toulouse School of Economics), the 2023 Barcelona Workshop in Financial Econometrics, the IAAE 2023 Annual Conference (Norwegian Business School), the 6th Workshop on Goodness-Of-Fit, Change-Point and Related Problems, the 2023 Robust Econometric Methods in Financial Econometrics (and Beyond) workshop, as well as seminar participants at CREST and Princeton. A. Rahbek and G. Cavaliere gratefully acknowledge support from the Independent Research Fund Denmark (DFF Grant 7015-00028) and the Italian Ministry of University and Research (PRIN 2017 Grant 2017TA7TYC and PRIN 2020 Grant 2020B2AKFW). T. Mikosch's research is partially supported by the Independent Research Fund Denmark (DFF Grant 9040-00086B).

Appendix A. Proofs

Proof of Lemma 2.1

Observe that the ACD Eq. (1.2) can be formulated as a stochastic recurrence equation (SRE):

$$x_i = A_i\, x_{i-1} + B_i, \qquad i \in \mathbb{Z}, \tag{A.1}$$

with a sequence $(A_i, B_i) = (\omega, \alpha)\varepsilon_i$, $i \in \mathbb{Z}$, of random vectors with i.i.d. positive $\{\varepsilon_i\}$. Using the SRE representation, it follows that $\{x_i\}$ is strictly stationary geometrically ergodic if and only if $\mathbb{E}[\ln(\alpha\varepsilon)] < 0$ by Theorem 2.1.3 and Proposition 2.2.4 in Buraczewski et al. (2016), BDM henceforth. The power-law tail behavior $\mathbb{P}(x > z) \sim c_\kappa\, z^{-\kappa}$, where

$$c_\kappa = \frac{\mathbb{E}[(\omega + (\alpha\,\varepsilon)\,x)^\kappa - ((\alpha\,\varepsilon)x)^\kappa]}{\kappa\,\mathbb{E}[(\alpha\varepsilon)^\kappa\ \ln(\alpha\varepsilon)]}, \tag{A.2}$$

follows from Theorem 2.4.4 in BDM. □

Proof of Lemma 2.2

Noting that for the exponential case

$$1 = \mathbb{E}[(\alpha\varepsilon)^\kappa] = \alpha^\kappa \int_0^\infty x^\kappa \exp(-x)\,dx = \alpha^\kappa\, \Gamma(\kappa+1), \tag{A.3}$$

the results hold by Lemma 2.1. □

Proof of Lemma 2.3

Convergence a.s. for $\kappa > 1$. Since $\kappa > 1$ we have $\mu = \mathbb{E}[x] < \infty$. We follow the argument in Theorem 5.1 in Gut (2009). Since $\{x_i\}$ is ergodic $T_n/n = \sum_{i=1}^n x_i/n \overset{\text{a.s.}}{\to} \mu$, hence $\nu_T = N_T + 1 \overset{\text{a.s.}}{\to} \infty$ and $T_{\nu_T}/\nu_T \overset{\text{a.s.}}{\to} \mu$. But $T < T_{\nu_T} \le T + x_{\nu_T}$ and

$$0 < T_{\nu_T}/\nu_T - T/\nu_T \le x_{\nu_T}/\nu_T \overset{\text{a.s.}}{\to} 0,$$

hence $\nu_T/T \overset{\text{a.s.}}{\to} 1/\mu$.
Convergence in distribution for $\kappa \ge 2$. We start with $\kappa > 2$. We have

$$T^{-1/2}(T_{N_T} - \mu N_T) \le T^{-1/2}(T - \mu\, N_T) \le T^{-1/2}(T_{N_T} - \mu N_T) + T^{-1/2} x_{\nu_T}. \tag{A.4}$$

First, we prove that $T^{-1/2} x_{\nu_T} \overset{p}{\to} 0$. For $M, \delta > 0$ we have

$$\begin{aligned}\mathbb{P}(T^{-1/2} x_{\nu_T} > M) &\le \mathbb{P}(T^{-1/2} x_{\nu_T} > M, |\nu_T/T - 1/\mu| > \delta) + \mathbb{P}(T^{-1/2} x_{\nu_T} > M, |\nu_T/T - 1/\mu| \le \delta)\\ &= I_1 + I_2.\end{aligned}$$

But $I_1 \to 0$ as $T \to \infty$ for every $\delta > 0$ by virtue of the first part of the proof. On the other hand, by stationarity,

$$I_2 \le \mathbb{P}\left(\max_{T(1/\mu-\delta)\le s\le T(1/\mu+\delta)} x_s > T^{1/2}\, M\right) \le \mathbb{P}\left(T^{-1/2} \max_{s\le 3\delta T} x_s > M\right).$$

The right-hand side converges to zero since $T^{-1/\kappa}\max_{s\le T} x_s$ converges in distribution to a Fréchet distribution; see BDM, Theorem 3.1.1.

In view of (A.4) we thus proved that the distributional limits of $T^{-1/2}(T_{N_T} - \mu N_T)$ and $T^{-1/2}(T - \mu\, N_T)$ coincide if they exist. However, Theorem 3.3.1 in BDM yields $n^{-1/2}(T_n - \mu\, n) \overset{d}{\to} N(0, \sigma^2)$ as $n \to \infty$ with $\sigma^2 = \mathbb{E}[(1+T_\infty)^2 - T_\infty^2]\mathbb{V}[x]$ and $T_\infty = \sum_{i=1}^\infty \alpha^i (\prod_{j=1}^i \varepsilon_j)$.

In what follows, we will frequently abuse notation: when sums are involved and their index is not a natural number we understand these expressions as taken at their integer parts. Abusing notation, we then have as $T \to \infty$,

$$T^{-1/2}\left(T_{T/\mu} - T\right) = \mu^{-1/2}\,(T/\mu)^{-1/2} \sum_{i=1}^{T/\mu} \left(x_i - \mu\right) \overset{d}{\to} N\left(0, \sigma^2/\mu\right).$$

Then the CLT for $T^{-1/2}(T_{N_T} - \mu N_T)$ will follow if we can prove that for every $M > 0$,

$$I = \mathbb{P}\left(T^{-1/2}\left|(T_{N_T} - \mu N_T) - (T_{T/\mu} - T)\right| > M\right) \to 0.$$

We apply an Anscombe type argument; see Gut (2009). For given $M, \delta > 0$ we have

$$I \le \mathbb{P}(|N_T/T - 1/\mu| > \delta) + \mathbb{P}\left(T^{-1/2}\left|(T_{N_T} - \mu N_T) - (T_{T/\mu} - T)\right| > M, |N_T/T - 1/\mu| \le \delta\right) = I_3 + I_4.$$

As before, $I_3 \to 0$ as $T \to \infty$. On the other hand, by stationarity,

$$I_4 \le \mathbb{P}\left(\max_{T(1/\mu-\delta)\le s\le T(1/\mu+\delta)} \left|T_s - T_{T/\mu} - \mu(s-T/\mu)\right| > T^{1/2}M\right)$$
$$\le 2\,\mathbb{P}\left(\max_{u\le\delta} |T_{uT} - \mu\, u\, T| > T^{1/2}\, M\right) .$$

Since $\{x_i\}$ is geometrically ergodic we can apply the functional CLT with Brownian limit and the continuous mapping theorem; see Theorem 19.1 in Billingsley (1999). Then the right-hand side vanishes by first letting $T \to \infty$ and then $\delta \to 0$.

The case $\kappa = 2$ is similar but we have to replace the normalization $T^{1/2}$ by $C(T\ln T)^{1/2}$ for a suitable constant $C > 0$. The proof of $x_{\nu_T}/(T\ln T)^{1/2} \xrightarrow{P} 0$ follows in the same way since $n^{-1/2}\max_{t=1,\dots,n} x_t$ converges in distribution to a Fréchet distribution; see BDM, Theorem 3.1.1. A functional CLT with Brownian limit and normalization $(T\ \ln T)^{1/2}$ is given in Guivarc'h and Le Page (2008).

Convergence in distribution for $1 < \kappa < 2$. Similar to the case of $\kappa > 2$, the starting point is the inequalities,

$$T^{-1/\kappa}(T_{N_T} - \mu N_T) \le T^{-1/\kappa}(T - \mu N_T) \le T^{-1/\kappa}(T_{N_T} - \mu N_T) + T^{-1/\kappa} x_{\nu_T}. \tag{A.5}$$

We observe that for $M, \delta > 0$, by stationarity,

$$\mathbb{P}(x_{\nu_T} > T^{1/\kappa}\, M) \le \mathbb{P}(|\nu_T/T - \mu| > \delta) + \mathbb{P}(T^{-1/\kappa} x_{\nu_T} > M\, , |\nu_T/T - \mu| \le \delta)$$
$$\le o(1) + \mathbb{P}\left(T^{-1/\kappa} \max_{s\le 3\delta T} x_s > M\right) ,$$

and the right-hand side converges to zero by first letting $T \to \infty$ and then $\delta \to 0$. In the last step one uses the Fréchet convergence of $T^{-1/\kappa}\max_{i=1,\dots,T} x_i$.

Next we observe that by BDM, Theorem 3.3.4,

$$(c_\kappa n)^{-1/\kappa}(T_n - \mu\, n) \xrightarrow{d} \left(\mathbb{E}\left[(1+T_\infty)^\kappa - T_\infty^\kappa\right]\right)^{1/\kappa} \eta_\kappa ,$$

where c_κ is the constant in (A.2), T_∞ is defined in the lemma and η_κ is κ-stable with characteristic function

$$\varphi_{\eta_\kappa}(s) = \exp\left(-\int_0^\infty (\exp(isy) - 1 - isy\,\mathbb{I}(1<\kappa<2))\,\kappa\, y^{-\kappa-1} dy\right) , \qquad s \in \mathbb{R} . \tag{A.6}$$

It remains to show that for every $M > 0$.

$$J = \mathbb{P}\left(T^{-1/\kappa}\left|(T_{N_T} - \mu N_T) - (T_{T/\mu} - T)\right| > M\right) \to 0 .$$

We have for every $\delta > 0$,

$$J \le o(1) + \mathbb{P}\left(T^{-1/\kappa}\left|(T_{N_T} - \mu N_T) - (T_{T/\mu} - T)\right| > M\, , |N_T/T - 1/\mu| \le \delta\right) = o(1) + J_1 .$$

Abusing notation, we have

$$J_1 \le \mathbb{P}\left(\max_{T(1/\mu-\delta)\le s\le T(1/\mu+\delta)} \left|T_s - T_{T/\mu} - \mu\,(s - T/\mu)\right| > T^{1/\kappa} M\right) \le 2\mathbb{P}\left(\max_{u\le T\delta} |T_u - \mu u| > T^{1/\kappa} M\right)$$

On one hand, we observe that for fixed $\widetilde{\varepsilon} > 0$,

$$\mathbb{P}\left(\max_{u\le T\delta} |T_u - \mu u| > T^{1/\kappa} M, \max_{s\le T\delta} x_s > T^{1/\kappa}\,\widetilde{\varepsilon}\right) \le \delta T\,\mathbb{P}(x > \widetilde{\varepsilon} T^{1/\kappa}) \sim \text{const}\ \delta\widetilde{\varepsilon}^{-\kappa} ,\ T \to \infty.$$

The right-hand side converges to zero as $\delta \to 0$. Next we mimic the proof of Theorem 4.5.2 in BDM. Write

$$X_T = \overline{X}_T + \underline{X}_T, \quad \overline{X}_T = X_T f(T^{-1/\kappa} X_T), \qquad \overline{T}_n = \sum_{i=1}^n \overline{X}_i ,$$

with $f(x) \in [0,1]$ smooth, $\operatorname{supp} f \subset \{x : |x| \le \widetilde{\varepsilon}\}$, and $f(x) = 1$ for $|x| \le \widetilde{\varepsilon}/2$. Then

$$\mathbb{P}\left(\max_{u\le\delta T} |T_u - \mu u| > T^{1/\kappa}\, M, \max_{s\le\delta T} x_s \le T^{1/\kappa}\,\widetilde{\varepsilon}\right) \le \mathbb{P}\left(\max_{u\le\delta T} |\overline{T}_u - u\,\mathbb{E}[\overline{x}]| + \delta T\,\mathbb{E}[\underline{x}] > T^{1/\kappa} M\right)$$

By Karamata's theorem (see Bingham et al., 1987) for large T,

$$\delta\, T^{1-1/\kappa}\,\mathbb{E}[\underline{x}] \ge \text{const}\ \delta\widetilde{\varepsilon}^{1-\kappa}\,\frac{\mathbb{E}[x/(\widetilde{\varepsilon} T^{1/\kappa})\mathbb{I}(x > \widetilde{\varepsilon} T^{1/\kappa})]}{\mathbb{P}(x > \widetilde{\varepsilon} T^{1/\kappa})} \sim \text{const}\ \delta\widetilde{\varepsilon}^{1-\kappa} \to 0 \text{ as } \delta \to 0.$$

Therefore it is suffices to show that the following quantity vanishes by first letting $T \to \infty$ and then $\delta \to 0$:

$$Q = \mathbb{P}\left(\max_{u\le\delta T} |\overline{T}_u - u\,\mathbb{E}[\overline{x}]| > T^{1/\kappa}\, M\right) .$$

With $s(T) = \delta\, T^{1-\beta}$ for $\beta \in (0,1)$, (Here we assume without loss of generality that $s(T)$ is an integer).

$$Q \le \mathbb{P}\left(\max_{k=1,\dots,s(T)} |\overline{T}_{kT^\beta} - kT^\beta\mathbb{E}[\overline{x}]| > T^{1/\kappa} M\right)$$

$$
\begin{aligned}
&+\mathbb{P}\left(\max_{k=1,\dots,s(T)}\max_{u\in\{(k-1)T^{\beta}+1,\dots,kT^{\beta}\}}|(\overline{T}_u-\overline{T}_{(k-1)T^{\beta}})-(s-(k-1)T^{\beta})\mathbb{E}[\overline{x}]|>T^{1/\kappa}M\right)\\
&=Q_1+Q_2\,,
\end{aligned}
$$

ignoring the last incomplete block of indices as it does not contribute to the asymptotic theory. Observe that $T^{\beta-1/\kappa}\mathbb{E}[\overline{x}]\to 0$ as $T\to\infty$ for $\beta<1/\kappa$. Hence by stationarity and for large T,

$$
\begin{aligned}
Q_2 &\le s(T)\,\mathbb{P}\left(\max_{u\le T^{\beta}}|\overline{T}_u-u\,\mathbb{E}[\overline{x}]|>T^{1/\kappa}M\right)\le s(T)\mathbb{P}(\overline{T}_{T^{\beta}}>T^{1/\kappa}M/2)\\
&\le s(T)\mathbb{P}(\overline{T}_{T^{\beta}}-\mathbb{E}[\overline{T}_{T^{\beta}}]>T^{1/\kappa}M/3)\le \text{const}\;\delta T^{(1-(2/\kappa))(1-\beta)}\mathbb{V}[T^{-\beta/\kappa}\overline{T}_{T^{\beta}}].
\end{aligned}
$$

By the calculations on p. 211 in BDM, the variance on the right-hand side is bounded. Hence $Q_2\to 0$.

Now we turn to Q_1. For $k\le s(T)$ we have for $\lambda>0$,

$$
\mathbb{P}(|\overline{T}_{kT^{\beta}}-kT^{\beta}\mathbb{E}[\overline{x}]|>\lambda)\le\lambda^{-2}\mathbb{V}[(kT^{\beta})^{-1/\kappa}\overline{T}_{kT^{\beta}}]k^{2/\kappa}T^{2\beta/\kappa}\le\text{const}\;\lambda^{-2}k^{2/\kappa}T^{2\beta/\kappa}\,,
$$

where we again used the variance bounds on p. 211 in BDM. An application of Theorem 10.2 in Billingsley (1999) yields

$$
Q_1\le\text{const}\;M^{-2}\,T^{-2/\kappa}\,T^{2\beta/\kappa}\,(\delta T^{1-\beta})^{2/\kappa}=\text{const}\;M^{-2}\,\delta^{2/\kappa}\to 0\,,\qquad \delta\to 0\,.
$$

This finishes the proof in the case $\kappa\in(1,2)$.

Convergence a.s. and in distribution for $0<\kappa<1$. We know that $\{x_i\}$ is ergodic and positive. Therefore for every $M>0$

$$
\liminf_{n\to\infty}\frac{T_n}{n}\ge\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}x_i\mathbf{1}(x_i\le M)=\mathbb{E}[x_1\mathbf{1}(x_1\le M)]\quad\text{a.s.}
$$

Since $\mathbb{E}[x]=\infty$ we can let $M\to\infty$ to conclude that T_n/n a.s. converges to ∞. Moreover, $N_T\to\infty$ a.s. using the fact that $\{N_T>k\}=\{\sum_{i=1}^{k}x_i<T\}$. Hence, we also have T_{N_T}/N_T a.s. converges to ∞. Noting that

$$
\frac{T_{N_T}}{N_T}\le\frac{T}{N_T}\le\frac{T_{N_T+1}}{N_T+1}\,\frac{N_T+1}{N_T}
$$

where the left- and right-hand expressions converge to ∞, it follows that N_T/T a.s. converges to 0.

Next, using Theorem 3.3.4 in BDM, for $z>0$

$$
\begin{aligned}
\mathbb{P}(T^{-\kappa}N_T\le z)&=\mathbb{P}(T_{zT^{\kappa}}>T)=1-\mathbb{P}(T_{zT^{\kappa}}\le T)=1-\mathbb{P}\left((c_\kappa zT^{\kappa})^{-1/\kappa}T_{zT^{\kappa}}\le(c_\kappa z)^{-1/\kappa}\right)\\
&\to 1-\mathbb{P}\left(\left(\mathbb{E}\left[(1+T_\infty)^{\kappa}-T_\infty^{\kappa}\right]\right)^{1/\kappa}\eta_\kappa\le(c_\kappa z)^{-1/\kappa}\right)\\
&=\mathbb{P}\left(1/\left(c_\kappa\mathbb{E}[(1+T_\infty)^{\kappa}-T_\infty^{\kappa}]\eta_\kappa^{\kappa}\right)\le z\right)\,,
\end{aligned}
$$

where η_κ has characteristic function (A.6). □

Proof of Lemma 3.1

The results for the score hold by using Lemma 2.3(i) and establishing, $T^{-1/2}S_T=T^{-1/2}S_{(T/\mu)}+o_{\mathbb{P}}(1)$, where $S_{(v)}=\sum_{i=1}^{\lfloor v\rfloor}\xi_i$. To see that $T^{-1/2}S_T-T^{-1/2}S_{(T/\mu)}=o_{\mathbb{P}}(1)$, note that for every $M>0$ and $\delta\in(0,\mu^{-1})$,

$$
\begin{aligned}
\mathbb{P}\left(T^{-1/2}\left[S_T-S_{(T/\mu)}\right]>M\right)&=\mathbb{P}\left(T^{-1/2}\left[S_T-S_{(T/\mu)}\right]>M,\;|N_T/T-1/\mu|>\delta\right)\\
&\quad+\mathbb{P}\left(T^{-1/2}\left[S_T-S_{(T/\mu)}\right]>M,\;|N_T/T-1/\mu|\le\delta\right)\\
&=K_1+K_2
\end{aligned}
$$

Here, $K_1\le\mathbb{P}\left(|N_T/T-\mu|>\delta\right)\to 0$, while, by stationarity,

$$
\begin{aligned}
K_2&\le\mathbb{P}\left(T^{-1/2}\max_{T(1/\mu-\delta)\le s\le T(1/\mu+\delta)}|S_{(s)}-S_{(T/\mu)}|>M\right)\\
&\le 2\,\mathbb{P}\left(T^{-1/2}\max_{u\le T\delta}|S_{(u)}|>M\right)\to 2\,\mathbb{P}\left(\max_{s\le\delta}|B(s)|>M/2\right)\,,
\end{aligned}
$$

as $T\to\infty$, where B is a Brownian motion. The right-hand side converges to zero as $\delta\to 0$. The result for the score then holds as $T^{-1/2}S_{(T/\mu)}\overset{d}{\to}(\tau\Omega/\mu)^{1/2}\mathbf{Z}$ by standard application of a CLT for martingale differences.

Turning to the information, then by the ergodic theorem and as $N_T\overset{\text{a.s.}}{\to}\infty$ it follows that $N_T^{-1}I_T\overset{\text{a.s.}}{\to}\Omega$; see Embrechts et al. (1997, Lemma 2.5.3). On the other hand, we have by Lemma 2.3(i), $N_T/T\overset{\text{a.s.}}{\to}1/\mu$. Thus $T^{-1}I_T\overset{\text{a.s.}}{\to}\Omega/\mu$. □

Proof of Lemma 3.2

Write

$$I_T(\theta) = -\frac{\partial^2 L_T(\theta)}{\partial\theta\partial\theta'} = \sum_{i=1}^{N_T}\left(2\,\varepsilon_i\,\frac{\psi_i(\theta_0)}{\psi_i(\theta)} - 1\right)\mathbf{v}_i(\theta)\mathbf{v}_i(\theta)',$$

where $\mathbf{v}_i(\theta) = (1, x_{i-1})'/\psi_i(\theta)$. The summands constitute a strictly stationary ergodic sequence with values in the space $\mathbb{C}$ of continuous functions of θ in a neighborhood $\mathcal{N}(\theta_0)$ of θ_0 equipped with the uniform distance. It is not difficult to see that the sup-norm of these summands has finite expected value. Therefore the summands obey the ergodic theorem in $\mathbb{C}$; see Theorem 2.1 in Section 4.2 of Krengel (1985). Since $N_T \overset{\text{a.s.}}{\to} \infty$ we conclude that uniformly over $\theta \in \mathcal{N}(\theta_0)$,

$$N_T^{-1} I_T(\theta) \overset{\text{a.s.}}{\to} \mathbb{E}\left[\left(2\,\frac{\psi_1(\theta_0)}{\psi_1(\theta)} - 1\right)\mathbf{v}_1(\theta)\mathbf{v}_1(\theta)'\right].$$

Moreover, together with Lemma 2.3 (iv) we conclude that, uniformly over $\theta \in \mathcal{N}(\theta_0)$,

$$T^{-\kappa} I_T(\theta) = (N_T/T^{\kappa})(N_T^{-1} I_T(\theta)) \overset{d}{\to} W(\theta) = \lambda_\kappa\,\mathbb{E}\left[\left(2\,\frac{\psi_1(\theta_0)}{\psi_1(\theta)} - 1\right)\mathbf{v}_1(\theta)\mathbf{v}_1(\theta)'\right].$$

If $\theta - \theta_0 = O(T^{-\kappa/2})$ then we also have

$$\sup_\theta \left|T^{-\kappa}\left(I_T(\theta) - I_T(\theta_0)\right)\right| = \left(N_T/T^{\kappa}\right)\sup_\theta\left|N_T^{-1}(I_T(\theta) - I_T(\theta_0))\right| \to 0\,,$$

in probability. Thus we verified conditions C1 and C2 in Sweeting (1980) and, in turn, Theorem 1 applies, yielding as desired

$$\left(T^{-\kappa/2} S_T, T^{-\kappa} I_T\right) \overset{d}{\to} \left(\left(\lambda_\kappa \Omega\right)^{1/2}\mathbf{Z}, \lambda_\kappa \Omega\right)$$

for a bivariate standard Gaussian vector $\mathbf{Z}$ independent of λ_κ. □

Proof of Theorem 3.1

The result follows by applications of Lemmas 11 and 12 in (Kristensen and Rahbek, 2010). With the notation there, set $Q_T(\theta) = T^{-1} L_T(\theta)$, $U_T = 1$, and $v_T = T$; then conditions (i),(ii) and (iv) of Lemmas 11 and 12 in (Kristensen and Rahbek, 2010) hold by Lemma 3.1 above, as $\partial^2 Q_T(\theta_0)/\partial\theta\partial\theta' \overset{p}{\to} \Omega/\mu$, and $\partial Q_T(\theta_0)/\partial\theta \overset{d}{\to} (\tau\Omega/\mu)^{1/2}\mathbf{Z}$. Next, consider condition (iii) of Lemma 11 in Kristensen and Rahbek (2010); see also Jensen and Rahbek (2004). It follows that in a compact neighborhood $\mathcal{U}(\theta_0)$ of θ_0,

$$\begin{aligned}\sup_{\theta\in\mathcal{U}(\theta_0)}\left|\partial^3 Q_T(\theta)/\partial\alpha^3\right| &\le T^{-1}\sum_{i=1}^{N_T}\left[2\frac{x_i x_{i-1}^3}{\psi_i^4(\theta)} + 3\left(2\frac{x_i x_{i-1}^3}{\psi_i^4(\theta)} + \frac{x_{i-1}^3}{\psi_i^3(\theta)}\right)\right] \\ &\le \text{const}\, T^{-1}\sum_{i=1}^{N_T}\left(1+\varepsilon_i\right),\end{aligned} \tag{A.7}$$

and condition (iii) holds as N_T/T and $N_T^{-1}\sum_{i=1}^{N_T}\varepsilon_i$ are $O_{\mathbb{P}}(1)$. For the remaining third-order derivatives similar arguments apply. □

Proof of Theorem 3.2

Similar to the proof of Theorem 3.1, set $Q_T(\theta) = T^{-\kappa} L_T(\theta)$, $U_T = 1$, and $v_T = T^{\kappa}$. As there, conditions (i),(ii) and (iv) of Lemmas 11 and 12 in Kristensen and Rahbek (2010) hold by Lemma 3.1. Likewise as in (A.7),

$$\sup_{\theta\in\mathcal{U}(\theta_0)}\left|\partial^3 Q_T(\theta)/\partial\alpha^3\right| \le \text{const}\, T^{-\kappa}\sum_{i=1}^{N_T}\left(1+\varepsilon_i\right).$$

and condition (iii) holds since N_T/T^{κ} and $T^{-\kappa}\sum_{i=1}^{N_T}\varepsilon_i$ are $O_{\mathbb{P}}(1)$. □

Proof of Corollary 4.1

For $\kappa > 1$ the result follows from Theorem 3.1 by standard arguments. For $\kappa < 1$, the result holds by the proof of Theorem 1 and Corollary 1 in Sweeting (1980). □

Appendix B. The remainder term

We noticed in Section 3 that the likelihood function in Eq. (1.3) in addition to the observed durations $\{x_i\}_{i=1}^{N_T}$ in $[0, T]$ misses out on a term for $t_{N_T} < T$, i.e., the term containing the information about no events occurring in $(t_{N_T}, T]$ is ignored. Thus, strictly speaking, $\hat{\theta}_T$ is not the MLE. While it is common practice to ignore this likelihood contribution in the ACD literature (as standard ARCH software is typically applied for estimation), this term is usually included in the related point process literature; cf. Daley and Vere-Jones (2008). By using the point process representation of the ACD process, as originally noted in Engle and Russell (1998), it holds that the ACD conditional intensity $\lambda(t|\left\{x_i\right\}_{i=0}^{N_t}) = 1/\psi_{N_t+1}$ which implies that the remainder term missing in (1.3), $R_T(\theta)$ say, is given by

$$R_T(\theta) = -\frac{T-t_{N_T}}{\psi_{N_T+1}}. \tag{B.1}$$

We now establish that under geometric ergodicity of $\{x_i\}$, $R_T(\theta)$ is asymptotically negligible.

Lemma B.1. *Consider the remainder term $R_T(\theta)$ given by B.1. Then for $\left\{x_i\right\}$ strictly stationary and geometrically ergodic, we have $R_T(\theta)T^{-1/2} \xrightarrow{p} 0$ for $\kappa > 1$, and $R_T(\theta)T^{-\kappa/2} \xrightarrow{p} 0$ for $\kappa < 1$.*

Proof. With $v_T = N_T + 1$ we have $t_{N_T} < T < t_{v_T}$,

$$\max\left(t_{v_T} - T, T - t_{N_T}\right) < x_{v_T},$$

and $R_T(\theta) = (T - t_{N_T})/\psi_{v_T} \leq \varepsilon_{v_T}$. For $\kappa > 1$, $\lim_{T\to\infty} v_T/T = \lim_{T\to\infty} N_T/T = 1/\mu$ a.s. and

$$\begin{aligned}\mathbb{P}(\varepsilon_{v_T} > \sqrt{T}) &= \mathbb{P}(\varepsilon_{v_T} > \sqrt{T}, |v_T/T - 1/\mu| > \gamma) + \mathbb{P}(\varepsilon_{v_T} > \sqrt{T}, |v_T/T - 1/\mu| \leq \gamma)\\ &= I_1 + I_2\end{aligned}$$

for $\gamma > 0$. For every $\gamma > 0$, $I_1 \leq \mathbb{P}\left(|v_T/T - 1/\mu| > \gamma\right) \to 0$, while for small γ,

$$\begin{aligned}I_2 &\leq \mathbb{P}\left(\max_{(1/\mu-\gamma)\leq s\leq T(1/\mu+\gamma)} \varepsilon_s > \sqrt{T}\right) \leq \mathbb{P}\left(\max_{s\leq 3\gamma T} \varepsilon_s > \sqrt{T}\right)\\ &\leq 3\gamma T\,\mathbb{P}(\varepsilon > \sqrt{T}) \to 0\,.\end{aligned}$$

Next, for $\kappa < 1$, v_T/T^κ and N_T/T^κ converge in distribution to λ_κ, and hence

$$\mathbb{P}(\varepsilon_{v_T} > T^{\kappa/2}) = \mathbb{P}(\varepsilon_{v_T} > T^{\kappa/2}, v_T/T^\kappa > M) + \mathbb{P}(\varepsilon_{v_T} > T^{\kappa/2}, v_T/T^\kappa \leq M) = K_1 + K_2.$$

Here $K_1 \leq \mathbb{P}\left(v_T/T^\kappa > M\right)$ is arbitrarily small for large M as $T \to \infty$ while

$$K_2 \leq \mathbb{P}\Big(\max_{s\in 3MT^\kappa} \varepsilon_s > T^{\kappa/2}\Big) \leq 3\,M\,T^\kappa\,\mathbb{P}(\varepsilon > T^{\kappa/2}) \leq \text{const}\,T^\kappa\,\mathbb{P}(\varepsilon > T^{\kappa/2}) \to 0\,,$$

as desired. □

References

Allen, D., Felix, C., McAleer, M., Peiris, S., 2008. Finite sample properties of the QMLE for the log-ACD model: Application to Australian Stocks. J. Econometrics 147, 163–185.
Aquilina, M., Budish, E., O'Neill, P., 2022. Quantifying the high-frequency trading Arms Race. Q. J. Econ. 137, 493–564.
Billingsley, P., 1999. Convergence of Probability Measures. Wiley, NY.
Bingham, N., Goldie, C., Teugels, J., 1987. Regular Variation. Cambridge University Press, Cambridge, UK.
Bollerslev, T., Li, J., Patton, J., Quaedvlieg, R., 2020. Realized semicovariances. Econometrica 88, 1515–1551.
Bollerslev, T., Medeiros, M.C., Patton, J., Quaedvlieg, R., 2022a. From zero to hero: realized partial (co)variances. J. Econometrics 231, 348–360.
Bollerslev, T., Patton, J., Quaedvlieg, R., 2022b. Realized semibetas: disentangling "good" and "bad" downside risks. J. Financ. Econ. 144, 227–246.
Buraczewski, D., Damek, E., Mikosch, T., 2016. Stochastic Models with Power-Law Tails. Springer, NY.
Carrasco, M., Chen, X. and, 2002. Mixing and moment properties of various GARCH and stochastic volatility models. Econom. Theory 18, 17–39.
Cavaliere, G., Lu, Y., Rahbek, A., Stærk-Østergaard, J., 2023. Bootstrap inference for Hawkes and general point processes. J. Econometrics 235, 133–165.
Daley, D.J., Vere-Jones, D., 2008. An Introduction to the Theory of Point Processes, Volume II: General Theory and Structure. Springer, Berlin.
Drees, H., Haan, L., Resnick, S., 2000. How to make a hill plot. Ann. Statist. 28, 254–274.
Embrechts, P., Klüppelberg, C., Mikosch, T., 1997. Modelling Extremal Events: For Insurance and Finance. Springer, Heidelberg.
Embrechts, P., Liniger, T., Lin, L., 2011. Multivariate Hawkes processes: an application to financial data. J. Appl. Probab. 48, 367–378.
Engle, R.F., 2000. The econometrics of ultra-high-frequency data. Econometrica 68, 1–22.
Engle, R.F., Russell, J.R., 1998. Autoregressive conditional duration: a new model for irregularly spaced transaction data. Econometrica 66, 1127–1162.
Fernandes, M., Grammig, J., 2006. A family of autoregressive conditional duration models. J. Econometrics 130, 1–23.
Fernandes, M., Medeiros, M.C., Veiga, A., 2016. The (semi-)parametric functional coefficient autoregressive conditional duration model. Econometric Rev. 35, 1221–1250.
Francq, C., Zakoïan, J.M., 2022. Testing the existence of moments for GARCH processes. J. Econometrics 227, 47–64.
Grammig, J., Maurer, K.-O., 2000. Non-monotonic hazard functions and the autoregressive conditional duration model. Econom. J. 3, 16–38.
Guivarc'h, Y., Le Page, E., 2008. On spectral properties of a family of transfer operators and convergence to stable laws for affine random walks. Ergodic Theory Dynam. Systems 28, 423–446.
Gut, A., 2009. Stopped Random Walks: Limit Theorems and Applications. Springer, NY.

Hamilton, J.D., Jordà, Ò., 2002. A model of the federal funds rate target. J. Polit. Econ. 110, 1135–1167.
Hautsch, N., 2012. Econometrics of Financial High-Frequency Data. Springer, Berlin.
Janson, S., 1983. Renewal theory for m-dependent variables. Ann. Probab. 11, 558–568.
Jensen, S.T., Rahbek, A., 2004. Asymptotic normality of the QMLE estimator of ARCH in the nonstationary case. Econometrica 72, 641–646.
Johansen, S., 1991. Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. Econometrica 59, 1551–1580.
Krengel, U., 1985. Ergodic Theorems. De Gryuter, Berlin.
Kristensen, D., Rahbek, A., 2010. Likelihood-based inference for cointegration with nonlinear error-correction. J. Econometrics 158, 78–94.
Lee, S., Hansen, B., 1994. Asymptotic theory for the GARCH$(1,1)$ quasi-maximum likelihood estimator. Econom. Theory 10, 29–52.
Li, Y., Mykland, P.A., Renault, E., Zhang, L., Zheng, X., 2014. Realized volatility when sampling times are possibly endogenous. Econom. Theory 30, 580–605.
Mokkadem, A., 1990. Propriétés de mélange des processus autorégressifs polynomiaux. Ann. L'I.H.P. Probab. Et Stat. 26, 219–260.
Pacurar, M., 2008. Autoregressive conditional duration models in finance: a survey of the theoretical and empirical literature. J. Econ. Surv. 22, 711–751.
Phillips, P.C.B., 1991. Optimal inference in cointegrated systems. Econometrica 59, 283–306.
Sin, C., 2014. QMLE of a standard exponential ACD model: Asymptotic distribution and residual correlation. Ann. Financial Econ. 09 (2).
Sweeting, T.J., 1980. Uniform asymptotic normality of the maximum likelihood estimator. Ann. Statist. 8, 1375–1381.
Trapani, L., 2016. Testing for (in)finite moments. J. Econometrics 191, 57–68.
Zolotarev, V.M., 1986. One-Dimensional Stable Distributions. American Mathematical Society, Providence, Rhode Island.