Convergence results for tractable inference in α-stable stochastic processes

Marina Riabiz, Tohid Ardeshiri and Simon Godsill
Department of Engineering, University of Cambridge, Cambridge, CB2 1PZ, UK
Emails: mr622@cam.ac.uk, ta417@cam.ac.uk, sjg30@cam.ac.uk

***Abstract*—The α-stable distribution is highly intractable for inference because of the lack of a closed form density function in the general case. However, it is well-established that the α-stable distribution admits a Poisson series representation (PSR) in which the terms of the series are a function of the arrival times of a unit rate Poisson process. In our previous work, we have shown how to carry out inference for regression models using this series representation, which leads to a very convenient conditionally Gaussian framework, amenable to tractable Gaussian inference procedures. The PSR has to be truncated to a finite number of terms for practical purposes. The residual error terms have been approximated in our previous work by a Gaussian distribution, and we have recently shown that this approximation can be justified through a Central Limit Theorem (CLT). In this paper we present a new and exact characterisation of the first and second moments of the residual series over finite time intervals for the unit rate Poisson process, correcting a previous version that was only true in the infinite time limit. This enables us to test through simulation the rapid convergence of the residual terms to a Gaussian distribution of the Poisson series residual. We test this convergence using both Q-Q plots and the classical Kolmogorov-Smirnov test of Gaussianity.**

I. Introduction

Among the best known and practically used results in the statistical analysis of time-series and other fields are the central limit theorems (CLTs). According to the classical CLT, the sample mean of independent identically distributed (iid) random variables with finite mean and variance converges in distribution to a Gaussian, when the number of terms goes to infinity. The requirement of identical distribution can be relaxed, while that of finite variance can be replaced with other conditions of finiteness, see for example [1]. The hypothesis of finite variance is restrictive for real world observations that exhibit extreme values more frequently than a Gaussian distribution would allow. Examples of such abrupt changes include variations presented by stock prices or insurance gains/losses in financial applications, and have been studied since the seminal works of [2] and [3]. Sudden changes are also studied in the climatological sciences, see for example [4] and [5]. Further applications can be found in various fields of engineering, such as communications and signal processing [6]–[9]. We refer to [10] for an extensive bibliography of application areas and existing works.

To cope with time series presenting extreme values, in this paper we consider discrete-time linear processes driven by non-Gaussian noise:[1]

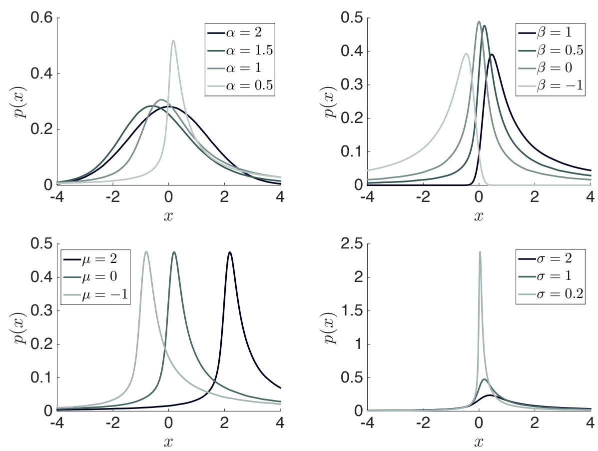

Fig. 1. Some α-stable pdfs $\mathcal{S}_\alpha(\sigma, \beta, \mu)$. If not specified $\alpha = 0.5$, $\sigma = 1$, $\beta = 0.5$, $\mu = 0$.

$$\mathbf{y} = \mathbf{G}\boldsymbol{\theta} + \mathbf{v}, \tag{1}$$

where $\mathbf{y}$ is a vector of N time series observations, $\mathbf{G}$ is a $N \times P$ matrix of regressors, $\boldsymbol{\theta}$ is a P-dimensional vector of unknown parameters and $\mathbf{v}$ is a N-dimensional vector of (typically iid) noise disturbances. When $\mathbf{v}$ is Gaussian distributed, simple and standard methods are available for Bayesian inference in such models, using closed form results combined with such methods as Variational Bayes or Monte Carlo methods. The aim of our approach is to show how these methods may be readily adapted to cases where the noise terms are general iid α-stable random variables, characterized by heavy tailed behaviour, but lacking closed form expression of the likelihood.

A. Contribution

In particular, in this work, we focus on the Poisson series representation (PSR) of the α-stable distribution, see [11]. The PSR was originally introduced by P. Lévy and formalised by [12]–[14]. The key result is that the sum of an infinite sequence of random variables (RVs), involving the arrival times of a Poisson process, converges almost surely (and hence in distribution) to an α-stable RV. For practical purposes,

The first author acknowledges support from EPSRC-DTG-2013/EU and The Alan Turing Institute under the EPSRC grant EP/N510129/1.

[1] A similar argument applies for nonlinear regressions where a solution can be obtained for the Gaussian case.

the full sequence cannot be generated, requiring truncation of the series; thus simulation and inference methods based on the PSR are approximate. In our recent work [15] we have proven that the residual series is asymptotically Gaussian, thus helping to justify the use of inference techniques based on conditionally Gaussian likelihoods. Furthermore, we have studied the theoretical convergence rate of the distribution of the residual to normality in the characteristic function domain for the general stable distribution, and we will report these new results in forthcoming publications.

In this paper we summarize these theoretical results on the PSR for α-stable random variables. We then provide experimental Gaussianity tests for the PSR residuals. We have also recently reported related results for multivariate stochastic integrals driven by α-stable Lévy processes in [16]. To perform these tests effectively, we derive non-asymptotic expressions of the moments of the residuals of the series. While the asymptotic moments are already stated in our earlier publications [15]–[21], we here provide for the first time an exact characterisation of the non-asymptotic variance of the residual, which corrects the previous versions which were only asymptotically true. We conclude the paper by summarizing a Bayesian inference scheme for the parameters $\boldsymbol{\theta}$ of our motivational example, the discrete time autoregressive model (1) driven by α-stable noise.

II. α-STABLE DISTRIBUTION

The α-stable distribution is of interest because of its versatility (capability to deal both with heavy-tailedness and skewness) and ease of interpretation through its parameters. It was originally introduced by [22] and it plays the key role of representing the limit distribution in a generalized version of the CLT, formalized by [23]. In this CLT the finite variance hypothesis of the classic CLT is relaxed, causing a power tail decay of the probability density function (pdf) of the form $p(x) \sim \frac{1}{|x|^{1+\alpha}}, |x| \to \infty$, where $\alpha \in (0,2)$ is the tail parameter. This asymptotic behaviour of the pdf corresponds to the presence of extreme values in the distribution, with more extreme values appearing more frequently for decreasing values of α. The other parameters of the distribution are $\beta \in [-1,1]$, that represents the skewness, $\mu \in (-\infty,\infty)$, that indicates the location and $\sigma > 0$, the scale. An α-stable distributed random variable X, $X \sim \mathcal{S}_\alpha(\sigma,\beta,\mu)$, has the following characteristic function (cf) $\phi(s)$

$$\log(\phi(s)) = \begin{cases} -\sigma^\alpha |s|^\alpha \left\{1 - i\beta \operatorname{sgn}(s) \tan \frac{\pi\alpha}{2}\right\} + i\mu s & \text{if } \alpha \neq 1, \\ -\sigma|s| \left\{1 + i\beta \operatorname{sgn}(s) \frac{2}{\pi} \log|s|\right\} + i\mu s & \text{if } \alpha = 1. \end{cases} \tag{2}$$

From (2) is possible to see that the Gaussian case is recovered for $\alpha = 2$, the Cauchy distribution for $\alpha = 1, \beta = 0$, and the Lévy distribution for $\alpha = 1/2, \beta = 1$.

Unlike the cf, the pdf of α-stable distributions is not available in closed form except in these few special cases. In Fig. 1 we give some pdf illustrations, produced by kernel smoothing histograms of samples generated through the exact sampling method of [24]. The lack of a closed form expression of the pdf complicates the inference in probabilistic models based on the α-stable distribution.

III. POISSON SERIES REPRESENTATION

If $X \sim \mathcal{S}_\alpha(\sigma,\beta,\mu)$, the PSR for α-stable RVs, as given in [11], states the following equality in distribution $\stackrel{\mathcal{D}}{=}$

$$X \stackrel{\mathcal{D}}{=} \sum_{j=1}^{\infty} W_j \Gamma_j^{-1/\alpha} - \mathbb{E}[W_1] b_j^{(\alpha)}, \tag{3}$$

where $\mathbb{E}[\cdot]$ denotes the expected value, $\{\Gamma_j\}_{j=1}^\infty$ are the arrival times of a unit rate Poisson process, and the $\{W_j\}_{j=1}^\infty$ are independent and identically distributed (i.i.d.) random variables independent of $\{\Gamma_j\}_{j=1}^\infty$, with $\mathbb{E}[|W_1|^\alpha] < \infty$. The coefficients $b_j^{(\alpha)}$ are non-zero only if $\alpha \in [1,2)$ and, for $\alpha \in (1,2)$ they have the telescoping structure

$$b_j^{(\alpha)} = \frac{\alpha}{\alpha - 1}\left(j^{\frac{\alpha-1}{\alpha}} - (j-1)^{\frac{\alpha-1}{\alpha}}\right).$$

From the PSR (3) it follows that, if $W_j \sim \mathcal{N}(\mu_W, \sigma_W^2)$, conditionally on the full sequence of arrival times $\{\Gamma_j\}_{j=1}^\infty$, X has Gaussian distribution

$$X \big| \{\Gamma_j\}_{j=1}^\infty \sim \mathcal{N}\left(\mu_W \sum_{j=1}^{\infty} \Gamma_j^{-1/\alpha} - b_j^{(\alpha)}, \sigma_W^2 \sum_{j=1}^{\infty} \Gamma_j^{-2/\alpha}\right). \tag{4}$$

As already noted, a serious issue is that the series (3) needs to be truncated, because an infinite sequence $\{\Gamma_j\}_{j=1}^\infty$ cannot be generated in practice. When only a truncated set of $\Gamma_j < c$ is known, where c is a truncation constant, the distribution of the first part of the series on the right hand side of (3), i.e. $\sum_{j:\Gamma_j \in [0,c]} W_j \Gamma_j^{-1/\alpha}$, is Gaussian, but the unconditional distribution of the residual term (as defined below) is not Gaussian. Note however that the heavy-tailed behaviour of the α-stable distribution is largely determined by the first part of the series, see for example [11], and so we can reasonably expect a CLT to apply to the residual series.

A. Asymptotic normality of the PSR residual

We here summarise the CLT required for justification of our conditionally Gaussian inference scheme. A more specialised result that did not account for the W_j terms in the residual was presented in [19]. The PSR can be split in terms of the $\{\Gamma_j \leq c\}$ as

$$X \stackrel{\mathcal{D}}{=} \sum_{j:\Gamma_j \in [0,c]} W_j \Gamma_j^{-1/\alpha} + R_{(c,\infty)},$$

where $R_{(c,\infty)}$ is the residual term, defined as $R_{(c,\infty)} := \lim_{d\to\infty} R_{(c,d)}$ where

$$R_{(c,d)} := \sum_{j:\Gamma_j \in (c,d)} W_j \Gamma_j^{-1/\alpha} - \mathbb{E}[W_1] \sum_{j=1}^{\lfloor d \rfloor} b_j^{(\alpha)}, \tag{5}$$

and $\lfloor \cdot \rfloor$ denotes the lower integer part. In [15] we have proved $R_{(c,\infty)}$ is asymptotically Gaussian, as $c \to \infty$, according to the following theorem.

Theorem 1: Assume $R_{(c,d)}$ as above and let $m_{(c,d)} := \mathbb{E}[R_{(c,d)}]$ denote its mean and $S^2_{(c,d)} := \mathbb{V}[R_{(c,d)}]$ its variance. If

$$\frac{\mathbb{E}[W_1^k]}{\mathbb{E}[W_1^2]^{k/2}} \frac{\alpha^{1-k/2}}{k!} \frac{(2-\alpha)^{k/2}}{k-\alpha} < \infty, \quad \forall k \geq 3,$$

then the following convergence in distribution holds, for $d \to \infty$, $c \to \infty$, $d \gg c$

$$Z_{(c,d)} := \frac{R_{(c,d)} - m_{(c,d)}}{S_{(c,d)}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1). \tag{6}$$

The limiting mean and variance of $R_{(c,\infty)}$, $m_{(c,\infty)}$ and $S^2_{(c,\infty)}$, have been characterized in [15]–[21] as

$$m_{(c,\infty)} := \mathbb{E}[R_{(c,\infty)}] = \lim_{d\to\infty} m_{(c,d)},$$
$$S^2_{(c,\infty)} := \mathrm{Var}[R_{(c,\infty)}] = \lim_{d\to\infty} S^2_{(c,d)}.$$

According to Theorem 1, these lead to the approximation

$$R_{(c,\infty)} \overset{\text{approx}}{\sim} \mathcal{N}\big(m_{(c,\infty)}, S^2_{(c,\infty)}\big),$$

for a sufficiently large value of c. Observe that Theorem 1 does not rely on the distribution of the W_j being Gaussian. However, if W_j are specified as Gaussian, we may approximate (4) with the overall conditionally Gaussian distribution

$$X|\{\Gamma_j \in [0,c]\} \overset{\text{approx}}{\sim} \mathcal{N}\left(\mu_W \sum_{j:\Gamma_j\in[0,c]} \Gamma_j^{-1/\alpha} + m_{(c,\infty)},\ \sigma_W^2 \sum_{j:\Gamma_j\in[0,c]} \Gamma_j^{-2/\alpha} + S^2_{(c,\infty)}\right), \tag{7}$$

which is our preferred framework for the use of the PSR in inference tasks.

IV. Non-asymptotic moments of the PSR residual

In this section we present for the first time the non-asymptotic moments $m_{(c,d)}$ and $S^2_{(c,d)}$. Previous versions reported in the literature (explicitly in [20] appendix B, [21] appendix 9.7, [18] Chapter 5, pages 46, 49, 51-52) were only correct in the limit $d \to \infty$, but here we require the non-asymptotic versions for the finite limit simulations carried out later in the paper. The required result is given in the following lemma.

Lemma 2: If the W_j have mean and variance μ_W and σ_W^2, respectively, then the mean and variance of the residual $R_{(c,d)}$, with integer d, are given by:

$$m_{(c,d)} = \begin{cases} \mu_W \frac{\alpha}{\alpha-1}\left(d^{\frac{\alpha-1}{\alpha}} - c^{\frac{\alpha-1}{\alpha}}\right) & \text{if } \alpha\in(0,1), \\ \mu_W \frac{\alpha}{\alpha-1}\left(-c^{\frac{\alpha-1}{\alpha}}\right) & \text{if } \alpha\in(1,2), \end{cases}$$

and

$$S^2_{(c,d)} = (\mu_W^2 + \sigma_W^2)\frac{\alpha}{\alpha-2}\left(d^{\frac{\alpha-2}{\alpha}} - c^{\frac{\alpha-2}{\alpha}}\right).$$

The derivation scheme requires an initial conditioning with respect to M, the random Poisson number of terms in $R_{(c,d)}$, followed by marginalization over M. We define the range of the truncated summation as $\mathcal{S} := \{j : \Gamma_j \in (c,d)\}$, and the random number of terms is $M := |\mathcal{S}|$. Note that by the properties of Poisson processes M is a Poisson random variable with mean equal to $d-c$. Moreover, we refer to the case $\alpha \in (0,1)$, noting that, when $\alpha \in (1,2)$, the expression of $R_{(c,d)}$ differs only for the constant $\mu_W \frac{\alpha}{\alpha-1} d^{\frac{\alpha-1}{\alpha}}$, given by the telescoping sum of the coefficients $b_j^{(\alpha)}$. This is simply subtracted from the expression of the mean, while the variance is not affected. We can now proceed to derive the required moments:

$$\begin{aligned} m_{(c,d)} &= \mathbb{E}\left[\mathbb{E}\left[\sum_{j\in\mathcal{S}} W_j \Gamma_j^{-1/\alpha}\middle| M\right]\right] \\ &= \mathbb{E}\left[\sum_{j\in\mathcal{S}} \mathbb{E}\left[W_j\Gamma_j^{-1/\alpha}|M\right]\right] \\ &= \mathbb{E}[M]\,\mathbb{E}\left[W_1\Gamma_1^{-1/\alpha}\right] \\ &= (d-c)\mathbb{E}\left[W_1\Gamma_1^{-1/\alpha}\right], \end{aligned}$$

and

$$\begin{aligned} \mathbb{E}[R^2_{(c,d)}] =& \mathbb{E}\left[\mathbb{E}\left[R^2_{(c,d)}\middle|M\right]\right] \\ =& \mathbb{E}\left[\mathbb{E}\left[\sum_{j\in\mathcal{S}} W_j\Gamma_j^{-1/\alpha}\sum_{i\in\mathcal{S}} W_i\Gamma_i^{-1/\alpha}\middle| M\right]\right] \\ =& \mathbb{E}[M]\,\mathbb{E}\left[W_1^2\Gamma_1^{-2/\alpha}\right] + \mathbb{E}\left[(M^2-M)\right]\left(\mathbb{E}\left[W_1\Gamma_1^{-1/\alpha}\right]\right)^2 \\ =& (d-c)\mathbb{E}\left[W_1^2\Gamma_1^{-2/\alpha}\right] + m^2_{(c,d)}, \end{aligned}$$

leading to

$$\begin{aligned} S^2_{(c,d)} =& \mathbb{E}[R^2_{(c,d)}] - (m_{(c,d)})^2 \\ =& (d-c)\mathbb{E}\left[W_1^2\Gamma_1^{-2/\alpha}\right]. \end{aligned} \tag{8}$$

We have also that, for $k = 1, 2$,

$$\begin{aligned} \mathbb{E}[\Gamma_i^{-k/\alpha}] &= \frac{1}{d-c}\int_c^d \Gamma^{-k/\alpha} d\Gamma = \frac{1}{d-c}\left[\frac{\alpha}{\alpha-1}\Gamma^{-k/\alpha+1}\right]_c^d \\ &= \frac{1}{d-c}\frac{\alpha}{\alpha-k}\left(d^{\frac{\alpha-k}{\alpha}} - c^{\frac{\alpha-k}{\alpha}}\right). \end{aligned}$$

Noting that the W_i are independent of the Γ_i and substituting these expressions into the moment calculations leads directly to the lemma.

These non-asymptotic moments are used in the following section to perform Gaussianity tests on the standardized residual $Z_{(c,d)}$.

V. Analysis of the convergence rate

In practice, for inference tasks, we need to generate a set $\{\Gamma_j < c\}$ for each stable random variable, meaning that c determines the average computational cost for using the PSR. Hence it is desired to perform the truncation as soon as c is sufficiently large for the Gaussian approximation on $Z_{(c,d)}$, as defined in (6), to hold.

In [15] we have studied the convergence rate of $Z_{(c,d)}$ to the

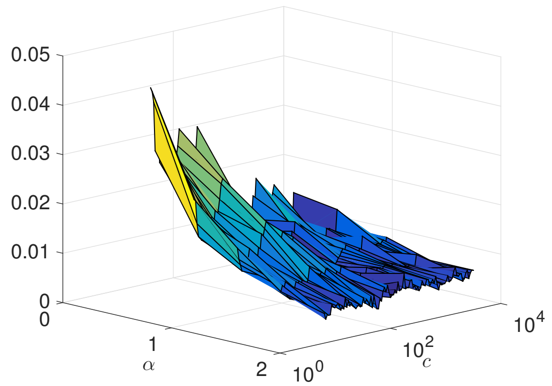

Fig. 2. Kolmogorov-Smirnov test statistic for $Z_{(c,d)}$, with $N = 10^4$, $d = 10^5, \mu_W = 1, \sigma_W = 1$, on a grid of c and α values.

standard Gaussian, with respect to c, in the characteristic function domain. In detail, we have shown that, as $\rho := d/c \to \infty$, the cf of $Z_{(c,d)}$ is

$$\phi_{Z_{(c,d)}}(s) = \exp\left(-s^2/2 + \xi\right),$$

where $\xi := \sum_{k=3}^{\infty} \overline{h}_k s^k$ and $\overline{h}_k$ are appropriately defined coefficients. As $c \to \infty$, ξ decays to zero, making $\phi_{Z_{(c,d)}}(s)$ tend to the characteristic function of a standard Gaussian. Furthermore, we have proven that this decay is exponential in c when $\mu_W = 0$, corresponding to the symmetric stable distribution with $\beta = 0$ (we refer to [11] for the non-linear transformations that map the moments of W_j and α into $\{\beta, \sigma\}$).

In this section we present an experimental study of the convergence rate in the pdf domain, for a generally skewed α-stable distribution. In particular, we perform the Kolmogorov-Smirnov test of Gaussianity for $Z_{(c,d)}$, see [25] for a description and further references. We recall here that the null hypothesis assumes that a given sample ($Z_{(c,d)}$ in our case) comes from a hypothesized distribution (central Gaussian here). The test statistic is then the supremum of the distance between the hypothesized cumulative density function (cdf) and the empirical cdf of the sample; a small value of the test statistic puts evidence in favour of the null hypothesis. We work with a $Z_{(c,d)}$ sample of size $N = 10^4$; we set $d = 10^5$ and we increase c on a logarithmic scale of values between 100 and 5×10^3. Furthermore we use $\mu_W = 1$, $\sigma_W = 1$, and we repeat for a grid of values of $\alpha \in \{0.3, 0.7, 1.1, 1.5, 1.9\}$, corresponding, respectively, to $\beta = \{0.75, 0.82, 0.86, 0.9, 0.92\}$, and $\sigma = \{1.52, 1.66, 1.90, 2.4, 4.8\}$. The result is reported in Fig. 2, showing the following two findings. Firstly, as we increase c for a fixed value of α, the test statistic decreases, supporting the hypothesis that the sample comes from a Gaussian distribution, as expected from Theorem 1. Secondly, for a fixed value of c, the test indicates that Gaussianity increases as $\alpha \to 2$; this is not surprising given that the α-stable distribution approaches the Gaussian, as $\alpha \to 2$.

To support the test, in Fig.3 we show Q-Q plots of the quantiles of the sample distribution of $Z_{(c,d)}$, compared to the standard Gaussian quantiles. We consider a 'very' heavy-tailed stable distribution ($\alpha = 0.5$) and a 'less' heavy tailed one ($\alpha = 1.8$), as well as two values for c, a lower one ($c = 20$) and a higher one ($c = 10^3$). As expected, the normalized residuals $Z_{(c,d)}$ are still heavy tailed when $\alpha = 0.5, c = 20$ and the Gaussianity is better met when c is increased. On the other hand, when $\alpha = 1.8$, the quantiles of the empirical distribution resemble the Gaussian already when $c = 20$.

VI. Inference in α-stable regression models

Here we summarize a possible Bayesian inference scheme for the parameters $\boldsymbol{\theta}$ of model (1), when $\mathbf{v} := [v_1, \ldots, v_N]$ has α-stable components. According to our approximation of the PSR, model (1) can be augmented to have a set of latent vectors, $\mathbb{T} := \{\boldsymbol{\Gamma}_n\}_{n=1}^N$, one for every element of $\mathbf{v}$. Each vector $\boldsymbol{\Gamma}_n := [\Gamma_{1,n}, \ldots, \Gamma_{M^n,n}]$ has different length, M^n, determined by finding the $\Gamma_{j,n}$ that are under the fixed threshold c. We compensate for the truncation by adding the residuals as in Theorem 1, and obtaining a conditionally Gaussian approximation for the distribution of each noise term $v_n, n = 1, \ldots, N$

$$v_n \big| \{\Gamma_{j,n} \in [0, c]\} \overset{\text{approx}}{\sim} \mathcal{N}\big(v_n \big| \mu_n, \sigma_n^2\big). \tag{9}$$

The mean μ_n and the variance σ_n^2 are expressed in the same way as in the right hand side term of equation (7). Denoting with $\boldsymbol{\mu} = [\mu_1, \ldots, \mu_N]'$ and with $\boldsymbol{\Sigma}$ the diagonal matrix with diagonal elements σ_n^2, $n = 1, \ldots, N$, (1) and (9) imply that the likelihood of data $\mathbf{y}$ can be written as

$$p(\mathbf{y}|\mathbf{G}\boldsymbol{\theta}, \mathbb{T}) = \mathcal{N}(\mathbf{y}|\mathbf{G}\boldsymbol{\theta} + \boldsymbol{\mu}, \boldsymbol{\Sigma}).$$

Regular inference can then be carried out as for the Gaussian model, by augmenting the set of parameters to be estimated to $\{\boldsymbol{\theta}, \mathbb{T}\}$. Assuming that the order of the model is known, a Gibbs sampler can be devised, that at the k-th iteration draws

$$\boldsymbol{\theta}^k \sim p(\boldsymbol{\theta}|\mathbb{T}^{k-1}, \mathbf{y}), \tag{10}$$

$$\mathbb{T}^k \sim p(\mathbb{T}|\boldsymbol{\theta}^k, \mathbf{y}). \tag{11}$$

For sampling $\boldsymbol{\theta}$ in (10), a conjugate Gaussian prior can be adopted, leading to a Gaussian conditional posterior with accordingly defined parameters. Regarding the step on $\mathbb{T}$, the posterior full conditional in (11) can be targeted with a Metropolis-within-Gibbs step. The posterior distribution of the parameters of interest $\boldsymbol{\theta}$ can be reconstructed by extracting the first component of the Markov chain run for N_{it} iterations $\{\boldsymbol{\theta}^k, \mathbb{T}^k\}_{k=1}^{N_{it}}$. We refer to [15], [17] for the details and a simulation result.

VII. Conclusions

In this work we have given an overview of a CLT for the residual of the PSR for α-stable random variables. This enables a conditionally Gaussian representation of the stable distribution. The latter is useful for inference in time series driven by α-stable noise, which currently is a research topic due to the lack of a closed-form probability density function. Furthermore, we have provided an experimental analysis of

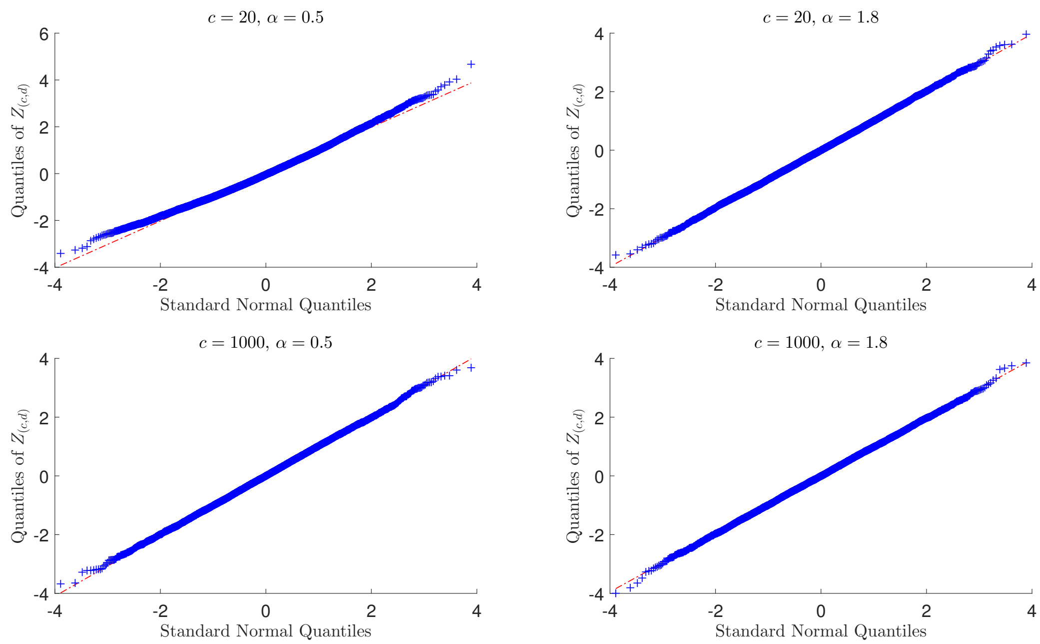

Fig. 3. Comparison of the quantiles of $Z_{(c,d)}$, with those of a standard Gaussian $d = 10^5, N = 10^4, \mu_W = 1, \sigma_W = 1$ for $c \in \{20, 1000\}$, $\alpha \in \{0.5, 1.8\}$.

the convergence rate of the empirical distribution of the PSR residual to the Gaussian distribution and characterized the non-asymptotic moments of the series residual.

References

[1] Lindeberg, J. W., "Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung," *Mathematische Zeitschrift*, vol. 15, pp. 211–225, 1922.

[2] Mandelbrot, B., "New methods in statistical economics," *Journal of Political Economy*, vol. 71, no. 5, pp. pp. 421–440, 1963.

[3] Fama, E. F., "The behavior of stock-market prices," *The Journal of Business*, vol. 38, no. 1, pp. pp. 34–105, 1965.

[4] Katz, R. W. and Brown, B. G., "Extreme events in a changing climate: Variability is more important than averages," *Climatic Change*, vol. 21, no. 3, pp. 289–302, 1992.

[5] Katz R. W. and Parlange M. B. and Naveau P., "Statistics of extremes in hydrology," *Advances in Water Resources*, vol. 25, no. 812, pp. 1287 – 1304, 2002.

[6] Nikias, C. L. and Shao, M., *Signal processing with alpha-stable distributions and applications*, ser. Adaptive and learning systems for signal processing, communications, and control. Wiley, 1995.

[7] Achim A. and Bezerianos A. and Tsakalides P. , "Novel bayesian multiscale method for speckle removal in medical ultrasound images," *IEEE Transactions on Medical Imaging*, vol. 20, no. 8, pp. 772–783, Aug 2001.

[8] Achim A. and Kuruoglu E. E. and Zerubia J., "Sar image filtering based on the heavy-tailed rayleigh model," *IEEE Transactions on Image Processing*, vol. 15, no. 9, pp. 2686–2693, 2006.

[9] Lombardi, M. J. and Godsill, S. J., "On-line Bayesian estimation of signals in symmetric α-stable noise," *Signal Processing, IEEE Transactions on*, vol. 54, no. 2, pp. 775–779, 2006.

[10] Nolan J. , April 2017. [Online]. Available: {http://fs2.american.edu/jpnolan/www/stable/StableBibliography.pdf}

[11] Samoradnitsky, G. and Taqqu, S., *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*, ser. Stochastic Modeling Series. Taylor & Francis, 1994.

[12] LePage, R. and Woodroofe, M. and Zinn, J., "Convergence to a Stable distribution via order statistics," *The Annals of Probability*, vol. 9, no. 4, pp. 624–632, 08 1981.

[13] LePage, R., "Appendix Multidimensional infinitely divisible variables and processes. Part I: Stable case," in *Probability theory on vector spaces IV*. Springer, 1989, pp. 153–163.

[14] ——, "Multidimensional infinitely divisible variables and processes Part II," in *Probability in Banach Spaces III*. Springer, 1981, pp. 279–284.

[15] Riabiz, M. and Ardeshiri, T. and Godsill, S. J., "A central limit theorem with application to inference in α-stable regression models," in *JMLR: Workshop and Conference Proceedings*, vol. 55 (NIPS 2016 Time Series Workshop Proceedings), February 2017.

[16] Riabiz, M. and Godsill, S.J., "Approximate simulation of linear continuous time models driven by asymmetric stable Lévy processes," in *Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on*, March 2017.

[17] Lemke, T. and Godsill, S.J., "Linear Gaussian computations for near-exact Bayesian Monte Carlo inference in skewed alpha-stable time series models," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, March 2012, pp. 3737–3740.

[18] Lemke, T., "Poisson series Approaches to Bayesian Monte Carlo Inference for Skewed α-Stable Distributions and Stochastic Processes," Ph.D. dissertation, 2014.

[19] Lemke, T. and Godsill, S. J., "A Poisson series approach to Bayesian Monte Carlo inference for skewed alpha-stable distributions," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 8023–8027.

[20] Lemke T. and Riabiz M. and Godsill S. J., "Fully Bayesian inference for α-stable distributions using a Poisson series representation," *Digital Signal Processing*, vol. 47, pp. 96 – 115, 2015, special Issue in Honour of William J. (Bill) Fitzgerald.

[21] Lemke, T., and Godsill. S. J., "Inference for models with asymmetric α-stable noise processes," in *Unobserved Components and Time Series Econometrics*, Koopman S. J. and Shephard N., Ed. Oxford: Oxford University Press, 2015, ch. 9.

[22] Lévy, P., *Calcul des Probabilités*, ser. PCMI collection. Gauthier-Villars, 1925.

[23] Gnedenko, B.V. and Kolmogorov, A.N., *Limit Distributions for Sums of Independent Random Variables*, ser. Addison-Wesley series in statistics. Addison-Wesley, 1968.

[24] Chambers, J. M. and Mallows, C. L. and Stuck, B. W., "A method for simulating Stable random variables," *Journal of the American Statistical Association*, vol. 71, no. 354, pp. 340–344, 1976.

[25] Massey F. J., "The Kolmogorov-Smirnov Test for Goodness of Fit," *Journal of the American Statistical Association*, vol. 46, no. 253, pp. 68–78, 1951.