

Simulated Convergence Rates with Application to an Intractable α -Stable Inference Problem

Marina Riabiz*, Tohid Ardestiri*†, Ioannis Kontoyiannis† 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

†Department of Electrical Engineering, Linköping University, Linköping, SE-581 83, Sweden

Email: tohid.ardeshiri@liu.se

‡Department of Informatics, Athens University of Economics and Business, Patission 76, Athens 10434, Greece

Email: yiannis@auer.gr

Abstract—We report the results of a series of numerical studies examining the convergence rate for some approximate representations of α -stable distributions, which are a highly intractable class of distributions for inference purposes. Our proposed representation turns the intractable inference for an infinite-dimensional series of parameters into an (approximately) conditionally Gaussian representation, to which standard inference procedures such as Expectation-Maximization (EM), Markov chain Monte Carlo (MCMC) and Particle Filtering can be readily applied. While we have previously proved the asymptotic convergence of this representation, here we study the rate of this convergence for finite values of a truncation parameter, c . This allows the selection of appropriate truncations for different parameter configurations and for the accuracy required for the model. The convergence is examined directly in terms of cumulative distribution functions and densities, through the application of the Berry theorems and Parseval theorems. Our results indicate that the behaviour of our representations is significantly superior to that of representations that simply truncate the series with no Gaussian residual term.

I. INTRODUCTION

The class of α -stable distributions is of interest because of its versatility (capability to deal both with heavy-tailedness and skewness) and ease of interpretation through its parameters. We refer to [1] for an extensive bibliography of existing works and application areas, ranging from engineering to finance to the climatological sciences. The distribution was originally introduced by [2] and it plays the key role of representing the limit distribution in a generalized version of the central limit theorem (CLT), formalized by [3]. 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 $p(x) \sim \frac{1}{|x|^{1+\alpha}}$, $|x| \rightarrow \infty$, 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 $\beta \in [-1, 1]$, the skewness, $\mu \in (-\infty, \infty)$, the location, and $\sigma > 0$, the scale. An α -stable distributed random variable X , $X \sim \mathcal{S}_\alpha(\sigma, \beta, \mu)$, has characteristic function (CF) $\phi_X(s) := \mathbb{E}[\exp(isX)]$ such that

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

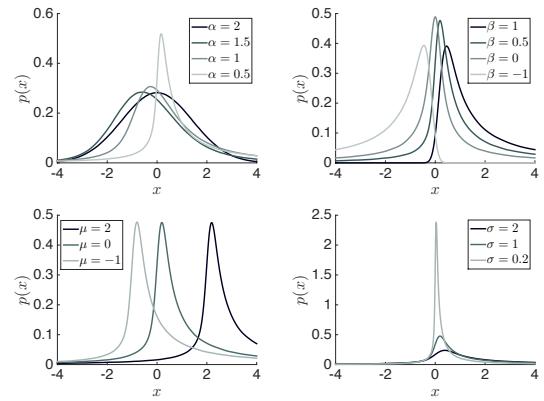


Fig. 1. Some α -stable pdfs $\mathcal{S}_\alpha(\sigma, \beta, \mu)$. If not specified, $\alpha = 0.5$, $\sigma = 1$, $\beta = 0.5$, $\mu = 0$.

From (1) it 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 $\alpha = 1/2, \beta = 1$. In contrast with the CF, the density function for 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 [4]. The lack of a closed form expression of the pdf complicates the inference in probabilistic models based on the α -stable distribution. Various approaches to inference have been proposed (see [5]–[12] for parameter inference) but most are approximate or unwieldy to implement. In [13]–[16] we have tackled the inference problem by resorting to very accurate approximations of the stable distribution via Poisson random series representations. We stress that the proposed approximation is relevant especially in large scale computation problems. In fact, as summarized in the following, we avoid the generation of an infinite series for each α -stable random variable by replacing the residual part of the series with a Gaussian random variable.

A. Poisson Sum Representation (PSR)

If we form the following infinite random summation,

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

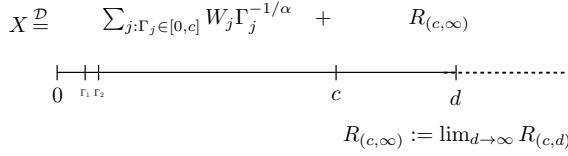


Fig. 2. Illustration of the PSR truncation in terms of the $\{\Gamma_j \leq c\}$.

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 $\mathbb{E}[|W_1|^\alpha] < \infty$, then the random variable X converges in distribution to the stable law $\mathcal{S}_\alpha(\sigma, \beta, 0)$, see [17, p.28]. The coefficients $b_j^{(\alpha)}$ are non-zero only if $\alpha \in [1, 2)$ and for this case they are readily computed and have a telescoping structure, see [17, p.28]. We further refer to [17, p.28] for the non-linear transformations that map the moments of W_j and α into the distribution parameters β and σ . We will not require the $b_j^{(\alpha)}$ coefficients or the non-linear transformations in this paper.

Now, the utility of the PSR for inference can be seen by choosing $W_j \sim \mathcal{N}(\mu_W, \sigma_W^2)$, the non-centered normal distribution. Then we can immediately write an auxiliary variables model for X as

$$X|\{\Gamma_j\}_{j=1}^\infty \sim \mathcal{N}\left(\mu_W m, \sigma_W^2 S^2\right),$$

$$m := \sum_{j=1}^\infty \Gamma_j^{-1/\alpha} - b_j^{(\alpha)}, \quad S^2 := \sum_{j=1}^\infty \Gamma_j^{-2/\alpha},$$

$$\Gamma_j - \Gamma_{j-1} \sim \mathcal{E}(1), \quad j = 1, 2, \dots, \infty, \quad \Gamma_0 = 0,$$

where $\mathcal{E}(1)$ is the exponential distribution with unity mean. In this model m and S^2 are treated as auxiliary random variables, and X has a conditionally Gaussian structure. This means that standard auxiliary variables methods for conditionally Gaussian models may be readily applied, for example blocked and collapsed Gibbs samplers [18], Rao-Blackwellised particle filters [19], [20] and Monte Carlo EM [21]. We note that the general framework here is a *scale* and *mean* mixture of normals, since we have:

$$p(X) = \int_{m \in \mathfrak{M}} \int_{S \in \mathfrak{S}^+} \mathcal{N}\left(X|\mu_W m, \sigma_W^2 S^2\right) p(m, S) dm dS.$$

For implementations of these ideas, see [13]–[16], [22].

While this exact representation of the stable law is a very appealing framework in theory, it is computationally intractable because of the infinite summations involved in m and S . A possible approach that we have adopted is to truncate the series to values of $\Gamma_j \leq c$, say, and to approximate the distribution of the residual term of the series. As illustrated in Figure 2, the PSR can be split as

$$X = X_c + R_{(c, \infty)}, \quad X_c := \sum_{j: \Gamma_j \in [0, c]} W_j \Gamma_j^{-1/\alpha}, \quad (2)$$

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

$\lim_{d \rightarrow \infty} R_{(c, d)}$, and

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

where $\lfloor \cdot \rfloor$ denotes the lower integer part.

It would seem natural that a CLT applies to $R_{(c, \infty)}$, and indeed the exact mean, $m_{(c, \infty)}$, and variance, $S_{(c, \infty)}^2$, of $R_{(c, \infty)}$ have been established along with its asymptotic normality (as $c \rightarrow \infty$) in [22] (see also [23] for a special case of the result when $\sigma_W = 0$). This CLT result helps to justify the Gaussian approximation $\hat{R}_{(c, \infty)}$ which is made in practical inference procedures:

$$\hat{R}_{(c, \infty)} \sim \mathcal{N}(m_{(c, \infty)}, S_{(c, \infty)}^2)$$

an approximation which converges to the distribution of the true residual $R_{(c, \infty)}$ as $c \rightarrow \infty$, and which is chosen to retain the conditionally Gaussian structure of the model:

$$X|\{\Gamma_j \in [0, c]\} \xrightarrow{\text{approx}} \mathcal{N}\left(\mu_W \sum_{j: \Gamma_j \in [0, c]} \Gamma_j^{-1/\alpha} + m_{(c, \infty)}, \sigma_W^2 \sum_{j: \Gamma_j \in [0, c]} \Gamma_j^{-2/\alpha} + S_{(c, \infty)}^2\right).$$

Denoting with $Z_{(c, \infty)} := (R_{(c, \infty)} - m_{(c, \infty)})/S_{(c, \infty)}$ the standardized residual, in [22] we showed that its CF, $\phi_{Z_{(c, \infty)}}(s)$, can be written as

$$\phi_{Z_{(c, \infty)}}(s) = \exp\left(-s^2/2 + \xi_c(s)\right), \quad (3)$$

where

$$\xi_c(s) := \sum_{k \geq 3} \frac{i^k}{k!} \frac{\frac{\alpha}{k-\alpha}}{\left(\frac{\alpha}{2-\alpha}\right)^{k/2}} c^{1-k/2} g_k s^k, \quad g_k := \frac{\mathbb{E}[W_1^k]}{\left(\mathbb{E}[W_1^2]\right)^{k/2}}. \quad (4)$$

Then the CLT for $R_{(c, \infty)}$ was proven showing that $\phi_{Z_{(c, \infty)}}(s)$ converges pointwise to the CF of the standard normal distribution $\mathcal{N}(0, 1)$

$$\phi(s) = \exp\left(-s^2/2\right). \quad (5)$$

B. Contributions of the paper

While the CLT proves that the approximation scheme becomes exact as the truncation limit c approaches infinity, we would ideally wish to have bounds on the error made for the finite values of c used in practice; hence the aims of this paper are to study numerically the error in distribution functions and density functions under finite values of c , potentially allowing automatic choice of the value of c for different parameter regimes. The contributions of the paper are to apply numerical integration methods to probability theoretic results in order to predict behaviour of the approximate model for finite c . We apply numerical integration to Berry theorems and Parseval theorems for a wide range of parameter settings, in order to test the Gaussianity of the residual series. We also provide preliminary results about the convergence of the whole approximated series to the α -stable law. We focus here on the symmetric stable law, in the expectation that this will give guidance to performance in general asymmetric stable laws. In

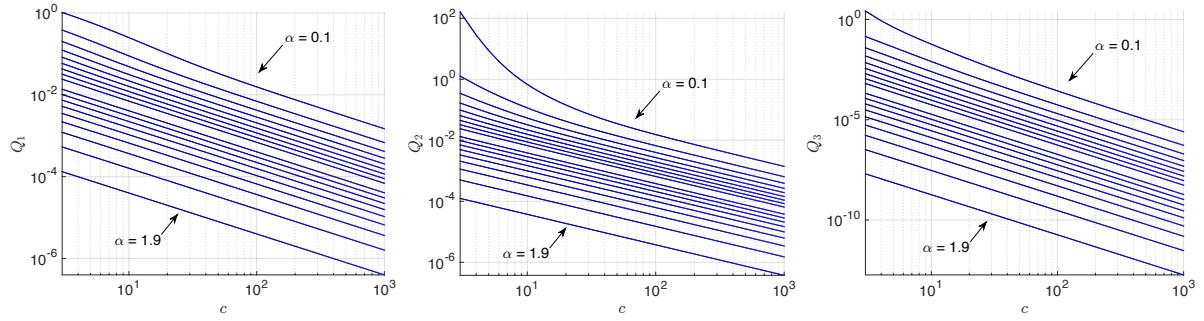


Fig. 3. Q_1, Q_2, Q_3 , the numerical approximations of I_1 (left), I_2 (centre) and I_3 (right), in a double logarithmic scale, as a function of $3 \leq c \leq 1000$ and $\alpha = 0.1, \dots, 1.9$, $\alpha \neq 1$, when $W_j \sim \mathcal{N}(0, \sigma_W^2)$. The plots include approximate error bands for each line.

our current work however, we are extending the frameworks to analyse convergence results in theory, without the need for numerical simulations, and to asymmetric stable laws.

II. DIVERGENCE MEASURES

Ideally, we would like to analytically convert the notion of convergence between CFs mentioned in section I-A to convergence between cumulative distribution functions (cdfs) and probability density functions (pdfs), the latter in terms of the commonly used divergence measures, such as Kullback-Leibler divergence (KLD) and α -divergence, see [24]. However, due to intractability of the stable pdf, such analytical study is not readily available in closed form. Hence, in the rest of this paper theoretical results which relate the deviation of CFs to deviations in cdfs and pdfs will be used in a numerical simulation to establish intuition for further research on this topic.

A. cdf divergence

If $F_{Z(c, \infty)}(x)$ is the cdf of $Z_{(c, \infty)}$ and $\Phi(x)$ is the cdf of the standard normal, then, by means of Berry's smoothing lemma, see [25, Lemma 2, p.538], for any $c \geq 1$, $\Omega > 0$ and $x \in \mathbb{R}$

$$\left| F_{Z(c, \infty)}(x) - \Phi(x) \right| \leq \frac{1}{\pi} \int_{-\Omega}^{\Omega} \frac{|\phi_{Z(c, \infty)}(s) - \phi(s)|}{|s|} ds + \frac{9.6}{\Omega} \quad (6)$$

$$\xrightarrow{\Omega \rightarrow \infty} \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{|\phi_{Z(c, \infty)}(s) - \phi(s)|}{|s|} ds := I_1, \quad (7)$$

where (7) is meaningful only if the integral converges, otherwise the truncated version (6) should be considered.

B. pdf divergence

Alternatively, if $\phi_{Z(c, \infty)}(s)$ and $\phi(s)$ are integrable functions, we can use the continuity theorem for densities, see [25, p.510]. Defining $f_{Z(c, \infty)}(x)$ and $f(x)$ to be the pdf of $Z_{(c, \infty)}$ and the standard normal pdf, respectively, we have

$$\left| f_{Z(c, \infty)}(x) - f(x) \right| \leq \frac{1}{2\pi} \int_{-\infty}^{\infty} \left| \phi_{Z(c, \infty)}(s) - \phi(s) \right| ds := I_2.$$

If $I_2 \xrightarrow{c \rightarrow \infty} 0$, this implies that the pdf of the residual converges uniformly to the standard normal pdf, given that I_2 does not depend on x .

C. Integrated squared error

Finally, we can consider convergence to zero of the integrated squared error (ISE) between pdfs, defined by

$$\text{ISE} \left(f_{Z(c, \infty)}(x), f(x) \right) := \int_{-\infty}^{\infty} \left(f_{Z(c, \infty)}(x) - f(x) \right)^2 dx.$$

If $\phi_{Z(c, \infty)}^2$ and $\phi^2(s)$ are integrable, then, using Parseval identity, see [25, p.510], we can express the ISE in terms of the CFs as follows

$$\text{ISE} \left(f_{Z(c, \infty)}(x), f(x) \right) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\phi_{Z(c, \infty)}(s) - \phi(s) \right)^2 ds := I_3.$$

III. NUMERICAL SIMULATIONS OF CONVERGENCE

In the following we assume that the integrability criteria in sections II-A to II-C are satisfied and we empirically verify them by checking that Q_1, Q_2, Q_3 , the numerical approximations of the integrals I_1, I_2, I_3 , do not diverge. In detail, we use the Gauss-Kronrod quadrature scheme throughout the MATLAB function 'quadgk' [26] in computing such approximations as a function of $c \geq 1$ and a range of different α values. We consider $W_j \sim \mathcal{N}(0, \sigma_W^2)$, corresponding to the symmetric stable distribution ($\beta = 0$). For this scenario, in [22] we have shown how the specific structure of the coefficients g_k in (4) can be exploited to provide a closed form expression of $\xi_c(s)$

$$\xi_c(s) = \psi_c(t) = c \left(1 - e^{-t} - t^a \gamma(1-a, t) + \frac{at}{(1-a)} \right),$$

with $a = \alpha/2$, $t = \frac{(1-a)s^2}{2ca}$ and $\gamma(s, t) := \int_0^t x^{s-1} e^{-x} dx$ the lower incomplete gamma function.

Figure 3 shows the dependence of Q_1, Q_2, Q_3 on c and α , on a double logarithmic scale, together with approximate numerical error bands. These do not appear visible, when compared to the Q_i values, suggesting empirically that the chosen numerical method is well behaved and has converged. The relation between $\log(Q_i)$, $i \in \{1, 2, 3\}$, and $\log(c)$ is very close to affine, noting that we might have to take $c \geq 10$ for small values of α . Then Figure 4 shows results of the linear regressions for finding u_i and v_i in

$$\log(Q_i) \approx v_i \log(c) + u_i, \quad i \in \{1, 2, 3\}, \quad c \geq 10,$$

corresponding to $Q_i \approx \exp(u_i) c^{v_i}$.

In detail, Q_1 and Q_3 have a decay approximately proportional to $1/c$, while Q_2 decays approximately as $1/c^2$, when $c \rightarrow \infty$.

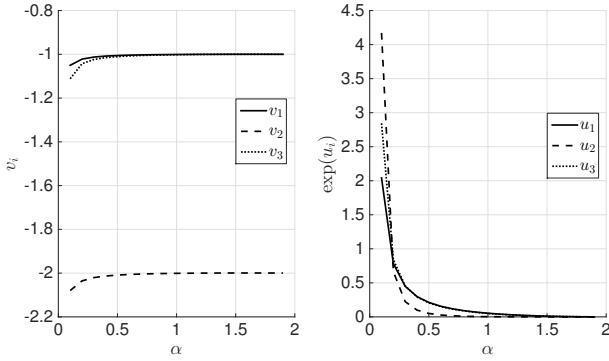


Fig. 4. Regression coefficients v_i and $\exp(u_i)$ as a function of α , when $W_j \sim \mathcal{N}(0, \sigma_W^2)$.

IV. APPROXIMATION ERROR IN THE STABLE LIKELIHOOD

In this section we examine the error caused by the overall approximation of the stable law with the PSR plus Gaussian residual. Ultimately this can be considered as the real error quantity of interest when performing likelihood or Bayesian inference with the model. Recalling that X_c is the truncated PSR, as in (2), define

$$X := X_c + R_{(c, \infty)}, \quad (\text{'true' stable law variate})$$

$$\hat{X} := X_c + \hat{R}_{(c, \infty)}, \quad (\text{Gaussian approximated series}).$$

Since $R_{(c, \infty)}$ and $\hat{R}_{(c, \infty)}$ are independent of X_c , by multiplicative property of the exponential (see e.g. [25, p.500]) we immediately have the following relations between CFs for these random variables

$$\phi_X(s) = \phi_{X_c}(s)\phi_{R_{(c, \infty)}}(s), \quad (8)$$

$$\phi_{\hat{X}}(s) = \phi_{X_c}(s)\phi_{\hat{R}_{(c, \infty)}}(s) = \phi_X(s)\phi_{\hat{R}_{(c, \infty)}}(s)/\phi_{R_{(c, \infty)}}(s). \quad (9)$$

We recall that $\phi_X(s)$ is given in closed form expression in equation (1), and the CFs for the true and approximate residuals are obtained as affine transformations of their normalised counterparts as (see e.g. [25, p. 499])

$$\phi_{R_{(c, \infty)}}(s) = \phi_{Z_{(c, \infty)}}(S_{(c, \infty)}s) \exp\left(ism_{(c, \infty)}\right),$$

$$\phi_{\hat{R}_{(c, \infty)}}(s) = \phi(S_{(c, \infty)}s) \exp\left(ism_{(c, \infty)}\right),$$

with $m_{(c, \infty)}$ and $S_{(c, \infty)}$ their mean and standard deviation, as in [22], $\phi_{Z_{(c, \infty)}}(s)$ defined in (3), and $\phi(s)$ defined in (5).

As in Section II-A, we aim to convert pointwise errors in the CF to convergence of the cdf. We then denote by $F_X(x)$, $F_{X_c}(x)$, $F_{\hat{X}}(x)$ the cdfs of X , X_c and \hat{X} , respectively. By means of the previous definitions in section I and eqs. (8) and (9) and using the same limiting technique as in Section II-A, we can bound the absolute difference of such cdfs as follows

$$|F_X(x) - F_{X_c}(x)| \leq \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{|\phi_X(s)|}{|s|} \left| 1 - \frac{1}{\phi_{R_{(c, \infty)}}(s)} \right| ds := I_1^{X_c},$$

$$|F_X(x) - F_{\hat{X}}(x)| \leq \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{|\phi_X(s)|}{|s|} \left| 1 - \frac{\phi_{\hat{R}_{(c, \infty)}}(s)}{\phi_{R_{(c, \infty)}}(s)} \right| ds := I_1^{\hat{X}}.$$

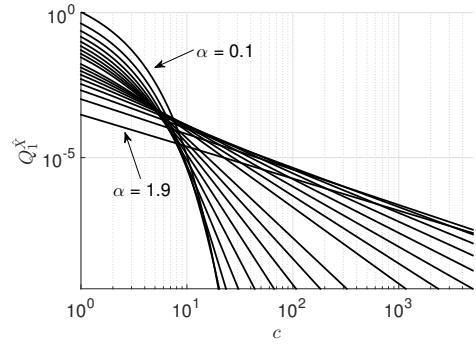


Fig. 5. Numerical approximation of $I_1^{\hat{X}}$, in a double logarithmic scale, as a function of $1 \leq c \leq 5000$ and $\alpha \in \{0.1, \dots, 1.9\}$, when $W_j \sim \mathcal{N}(0, 2)$.

Given that $|\phi_{X_c}(s)| < 1$, $I_1^{\hat{X}}$ is finite when I_1 is finite, while we have no guarantees on $I_1^{X_c}$. We have in fact proved analytically (proof to be presented in a later publication) that, when $\mu_W = 0$, $I_1^{X_c}$ is larger than $I_1^{\hat{X}}$ for any value of $\alpha \in (0, 2]$ and $c > \log(2)/(\gamma(1 - a, 1) + e^{-1} - 1)$, where $a = \alpha/2$. Therefore our Gaussian residual approximation is likely to have lower error in cdf than direct truncation of the series, a result borne out by previous experimental simulations, see [23, p.56-57]. Hence, in the simulations in Figure 5 we show just Q_1^X the numerical approximation of the smaller error term $I_1^{\hat{X}}$, noticing that the numerical approximation of $I_1^{X_c}$ does not seem to be finite when $\alpha > 1$. We observe that Q_1^X is asymptotically lower for smaller values of α , a reversal of the trend in the residual convergence to Gaussianity shown in Fig. 3. We believe this is because the residual term is of much less relative significance compared to the heavy-tailed initial terms in the PSR as α decreases. We can also observe that the rate of convergence is dramatically improved for smaller α , again in contrast with Fig. 3.

V. DISCUSSION AND CONCLUSION

In this paper we have generated numerical simulation results to characterise the error made when approximating the otherwise intractable PSR residual series as a Gaussian. The results show that convergence is overall poorer for smaller values of α , which seems reasonable since these cases are further from the Gaussian limit of the stable law ($\alpha \rightarrow 2$). Rate of convergence as c increases appears to be approximately $O(1/c)$ for distribution functions, almost independent of α . These trends appear to be somewhat reversed in our initial studies for the approximation error for the whole stable distribution (i.e. not just analysing the residual term) in that errors are smaller for smaller α . This is probably because the residual terms are dwarfed in size by the larger terms in the very heavy-tailed earlier parts of the series. In addition, our approximation of the residual is found to be significantly more accurate than the more traditionally analysed approach [27], [28] which simply truncates the series without approximating the residual. In our current work we have proved some theoretical convergence rates for the symmetric stable laws, and are extending these theoretical results to the asymmetric case. These further results will be presented in future publications. Many of our results can also be extended to the case of continuous-time α -stable Lévy processes, see [16], [29]–[32] and this is also a topic of current investigation.

REFERENCES

[1] Nolan, J. , June 2017. [Online]. Available: {<http://fs2.american.edu/jpnolan/www/stable/StableBibliography.pdf>}

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

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

[4] 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.

[5] McCulloch, J. H., “Simple consistent estimators of stable distribution parameters,” *Communications in Statistics - Simulation and Computation*, vol. 15, no. 4, pp. 1109–1136, 1986.

[6] Kuruoglu, E.E., “Density parameter estimation of skewed α -stable distributions,” *Signal Processing, IEEE Transactions on*, vol. 49, no. 10, pp. 2192–2201, Oct 2001.

[7] Koutrouvelis, I. A., “Regression-type estimation of the parameters of Stable laws,” *Journal of the American Statistical Association*, vol. 75, no. 372, pp. 918–928, 1980.

[8] Nolan, J., “Maximum likelihood estimation and diagnostics for stable parameters,” *Levy processes: Theory and applications*, pp. 379–400, 2001.

[9] Buckle, D. J., “Bayesian inference for Stable distributions,” *Journal of the American Statistical Association*, vol. 90, no. 430, pp. 605–613, 1995.

[10] Godsill, S., “Inference in symmetric alpha-stable noise using MCMC and the slice sampler,” in *Acoustics, Speech, and Signal Processing, 2000. ICASSP '00. Proceedings. 2000 IEEE International Conference on*, vol. 6, 2000, pp. 3806–3809 vol.6.

[11] Riabiz, M. and Lindsten F., and Godsill S., “Pseudo-Marginal MCMC for Parameter Estimation in α -Stable Distributions,” *IFAC-PapersOnLine*, vol. 48, no. 28, pp. 472 – 477, 2015, 17th IFAC Symposium on System Identification SYSID 2015.

[12] Lombardi, M. J., “Bayesian inference for α -stable distributions: A random walk {MCMC} approach ,” *Computational Statistics & Data Analysis*, vol. 51, no. 5, pp. 2688 – 2700, 2007.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

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

[18] Van Dyk, D. A. and Park, T., “Partially collapsed gibbs samplers: Theory and methods,” *Journal of the American Statistical Association*, vol. 103, no. 482, pp. 790–796, 2008.

[19] Schon, T. and Gustafsson, F. and Nordlund, P. J., “Marginalized particle filters for mixed linear/nonlinear state-space models,” *IEEE Transactions on Signal Processing*, vol. 53, no. 7, pp. 2279–2289, July 2005.

[20] Doucet, A. and de Freitas, N. and Murphy, K. P. and Russell, S. J., “Rao-Blackwellised Particle Filtering for Dynamic Bayesian Networks,” in *Proceedings of the 16th Conference on Uncertainty in Artificial Intelligence*, ser. UAI '00. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2000, pp. 176–183.

[21] Tanner, M.A., *Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions*, ser. Springer Series in Statistics. Springer New York, 1997.

[22] Riabiz, M. and Ardeshiri, T. and Godsill, S., “A central limit theorem with application to inference in α -stable regression models,” in *Proceedings of the Time Series Workshop at NIPS 2016*, ser. Proceedings of Machine Learning Research, vol. 55. Barcelona, Spain: PMLR, 09 Dec 2017, pp. 70–82.

[23] Lemke, T., “Poisson Series Approaches to Bayesian Monte Carlo Inference for Skewed α -Stable Distributions and Stochastic Processes,” Ph.D. dissertation, Kaiserslautern University of Technology, 2014.

[24] Minka, T., “Divergence measures and message passing,” Technical report, Microsoft Research, Tech. Rep., 2005.

[25] Feller, W., *An introduction to probability theory and its applications*, ser. Wiley mathematical statistics series. Wiley, 1966, no. v. 2.

[26] Shampine, L.F., “Vectorized adaptive quadrature in {MATLAB},” *Journal of Computational and Applied Mathematics*, vol. 211, no. 2, pp. 131 – 140, 2008.

[27] Paulauskas, V. and Ledoux, M., “A rate of convergence in the poissonian representation of stable distributions,” *Lithuanian Mathematical Journal*, vol. 36, no. 4, pp. 388–399, 1996.

[28] Bentkus V. and Gtze F. and Paulauskas V., “Bounds for the accuracy of poissonian approximations of stable laws,” *Stochastic Processes and their Applications*, vol. 65, no. 1, pp. 55 – 68, 1996.

[29] Asmussen, S. and Rosinski, J., “Approximations of small jumps of Lévy processes with a view towards simulation,” *Journal of Applied Probability*, vol. 38, no. 2, pp. 482–493, 2001.

[30] Rosinski, J., “Series representations of Lévy processes from the perspective of point processes,” in *Lévy Processes*, O. Barndorff-Nielsen, S. Resnick, and T. Mikosch, Eds. Birkhauser Boston, 2001, pp. 401–415.

[31] ——, *Simulation of Lévy Processes*. John Wiley & Sons, Ltd, 2014.

[32] Riabiz, M. and Godsill, S. J., “Approximate simulation of linear continuous time models driven by asymmetric stable Lévy processes,” in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 2017, pp. 4676–4680.