



**HAL**  
open science

# Parameter estimation for stable distributions and their mixture

Omar Hajjaji, Solym Mawaki Manou-Abi, Yousri Slaoui

► **To cite this version:**

Omar Hajjaji, Solym Mawaki Manou-Abi, Yousri Slaoui. Parameter estimation for stable distributions and their mixture. *Journal of Applied Statistics*, 2024, pp.1-34. 10.1080/02664763.2024.2434627 . hal-04218721v2

**HAL Id: hal-04218721**

**<https://hal.science/hal-04218721v2>**

Submitted on 26 Dec 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Public Domain

# Parameter estimation for stable distributions and their mixture

Omar Hajjaji<sup>1</sup>, Solym Mawaki Manou-Abi<sup>1,2,3\*</sup> and Yousri Slaoui<sup>1</sup>

<sup>1</sup>\*Laboratoire de Mathématiques et Applications. Université de Poitiers, UMR CNRS 7348, Futuroscope Chasseneuil, Poitiers, 86000, France.

<sup>2</sup>Institut Montpelliérain Alexander Grothendieck. Université de Montpellier, UMR CNRS 5149, Place Eugene Bataillon, Montpellier, 34090, France.

<sup>3</sup>Université de Mayotte, Rue de l'Université, 97660, Dombeni, 97660, Mayotte, France.

\*Corresponding author(s). E-mail(s):

[solym.manou.abi@univ-poitiers.fr](mailto:solym.manou.abi@univ-poitiers.fr);

Contributing authors: [omar.hajjaji@univ-poitiers.fr](mailto:omar.hajjaji@univ-poitiers.fr);

[yousri.slaoui@univ-poitiers.fr](mailto:yousri.slaoui@univ-poitiers.fr);

## Abstract

In this paper, we consider estimating the parameters of univariate  $\alpha$ -stable distributions and their mixtures. First, using a Gaussian kernel density distribution estimator, we propose an estimation method based on the characteristic function. The optimal bandwidth parameter was selected using a plug-in method. We highlight another estimation procedure for the Maximum Likelihood framework based on the False position algorithm to find a numerical root of the log-likelihood through the score functions. For mixtures of  $\alpha$ -stable distributions, the EM algorithm and the Bayesian estimation method have been modified to propose an efficient and valuable tool for parameter estimation. The proposed methods can be generalised to multiple mixtures, although we have limited the mixture study to two components. A simulation study is carried out to evaluate the performance of our methods, which are then applied to real data. Our results appear to accurately estimate mixtures of  $\alpha$ -stable distributions. Applications concern

2 *Parameter estimation for stable distributions and their mixtures*

the estimation of the number of replicates in the Mayotte COVID-19 dataset and the distribution of the N-acetyltransferase activity of the Bechtel et al. data for a urinary caffeine metabolite implicated in carcinogens. We compare the proposed methods together with a detailed discussion and conclude with some other forthcoming work.

**Keywords:** Stable distribution; Parametric estimation; Newton-Raphson algorithm; Bisection algorithm; Mixture model; EM algorithm; Gibbs sampling algorithm, Metropolis-Hastings algorithm.

## 1 Introduction

In recent decades, many researchers have shown an interest in studying  $\alpha$ -stable distributions because of their ability to generalise widely used laws such as Gaussian, Lévy and Cauchy to handle impulsive and skewed data, which is particularly important in the financial field. In 1925, Lévy discovered that  $\alpha$ -stable distributions arise as the limit of normalised sums of independent and identically distributed random variables. The family of  $\alpha$ -stable distributions has skewness and tail thickness. Unfortunately, there is no closed-form expression for the cumulative distribution function and the probability density function, except in a few cases such as the Gaussian, Lévy and Cauchy distributions. Let  $X$  be a  $\alpha$ -stable random variable and  $\phi(t) = E(\exp(itX))$  the characteristic function. Note that this family of laws has multiple parameterizations. We follow the presentation in [1–4] and consider two types of representations. A random variable  $X \sim S(\alpha, \beta, \gamma, \zeta; 0)$  if  $\phi(t)$  is expressed as follows

$$\phi(t) = \begin{cases} \exp(-\gamma^\alpha |t|^\alpha [1 + i\beta(\tan(\frac{\pi\alpha}{2}))\text{sign}(t)(|\gamma t|^{1-\alpha} - 1)] + i\zeta t) & \text{if } \alpha \neq 1. \\ \exp(-\gamma |t| [1 + i\beta \frac{2}{\pi} \text{sign}(t) \log(\gamma |t|)] + i\zeta t) & \text{if } \alpha = 1. \end{cases} \quad (1)$$

where  $\alpha \in (0, 2]$  is the index of stability that governs the heaviness of the tail,  $\beta \in [-1, 1]$  the skewness parameter,  $\gamma > 0$  the scale parameter,  $\zeta \in \mathbb{R}$  the location or shift parameter and the sign function defined by:

$$\text{sign}(t) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0. \end{cases}$$

A random variable  $X$  is  $S(\alpha, \beta, \gamma, \mu; 1)$ , [1] if,

$$\phi(t) = \begin{cases} \exp(-\gamma^\alpha |t|^\alpha [1 - i\beta(\tan(\frac{\pi\alpha}{2}))\text{sign}(t)] + i\mu t) & \text{if } \alpha \neq 1. \\ \exp(-\gamma |t| [1 + i\beta \frac{2}{\pi} \text{sign}(t) \log(|t|)] + i\mu t) & \text{if } \alpha = 1. \end{cases} \quad (2)$$

These parameterizations are therefore usually characterised by four parameters:  $\alpha \in [0, 2]$ ,  $\beta \in [-1, 1]$ ,  $\gamma > 0$  and  $\zeta$  or  $\mu$ . If  $\beta = 0$  we have a symmetric stable distribution and the two parameterizations above are identical. If  $\gamma = 1$  and  $\mu = 0$  (or  $\zeta = 0$ ), we say that  $X$  is a standard  $\alpha$ -stable random variable. The different parameterizations have repeatedly led to misunderstandings. The parameterization (2) proposed by [1] does not provide continuity of the density function at points  $\alpha = 1$  and  $\beta = 0$  (because of the  $\tan(\pi\alpha/2)$  term), nor does it provide a scale-location family at  $\alpha = 1$  (because of the  $\gamma \log(\gamma)$  term), while the parameterization (1) is continuous with respect to all parameters. The above two formulations are connected by the key equation

$$\zeta = \begin{cases} \mu + \beta\gamma \tan(\frac{\pi\alpha}{2}) & \text{if } \alpha \neq 1 \\ \mu + \beta\frac{2}{\pi}\gamma \log(\gamma) & \text{if } \alpha = 1. \end{cases}$$

It is worth noting that  $\alpha$ -stable distributions have infinite variance for any  $\alpha < 2$ , and the mean is infinite for  $\alpha \in (0, 1]$ , which makes estimating the parameters difficult. Note that the mean, if it exists ( $\alpha > 1$ ), is the natural measure of the location, but it cannot be estimated as precisely as  $\zeta$ . In the case of  $\alpha = 2$ , we get the Gaussian distribution, when  $\alpha = 1$  and  $\beta = 0$ , it becomes the Cauchy distribution, and Lévy for  $\alpha = \frac{1}{2}$  and  $\beta = 1$ . It is recommended to use the 0 parameterization for numerical work and statistical inference. The standard mean-focus 1 parameterization is suitable for studying algebraic properties, we refer the reader to the work of [5]. Simulation methods are available for  $\alpha$ -stable distributions, for example Kanter [6] was the first to generate  $\alpha$ -stable random variables with  $\alpha \in (0, 1)$ . Later, Chambers, Mallows and Stuck [7] extended the method to the general case. It is important to note that when  $\alpha$  is close to 1 or 0 with  $\beta \neq 0$ , the computations are typically much more numerically demanding and the results may not be very accurate, see `Matlab` [2] and `R` [8]. Since it is not known when exactly the numerical difficulties occur, we only warn the reader to pay great attention in their programs and practical purposes. In the `stabledist` package [8], the authors highlight and discuss such warnings. These warnings are not yet the subject of explicit studies, and in the future we plan to investigate the explicit numerical behaviour in the vicinity of these boundary points. The simulation of  $\alpha$ -stable multivariate distributions has been the subject of research in [9]. The performance of our simulations is carried out with values of  $\alpha$  that are not close to 1. In the context of real applications, our applications concern the case  $\alpha \in (1, 2)$ . For  $\beta = 1$ , the distributions are maximally skewed to the right, and the `FMStable` package [10] provides distribution functions that are faster and more accurate than the `stabledist` package.

There are several methods in the scientific literature to estimate the parameters of  $\alpha$ -stable distributions, for example the Fractional Lower Order Moments (FLOM) method [11]; the quantile method improved by McCulloch [12], the Empirical Characteristic Function (ECF) and Maximum Likelihood (ML) approximation method [2], although the lack of a closed-form expression for the probability density is a theoretical disadvantage. Finite

mixture models are becoming increasingly popular and play a crucial role in density estimation, and the mixture of  $\alpha$ -stable distributions is a popular tool for modelling skewed and impulsive data, making it applicable in various fields. Two well-known methods of mixture estimation are the Expectation-Maximisation (EM) algorithm and the Bayesian approach. The EM algorithm is particularly useful in cases where the data are incomplete or partially observed. Bayesian estimation uses prior knowledge of the probability distribution of the parameters to estimate the value of the unknown parameters. Much work has been done in this direction, including the papers cited by [13] and references therein. Let us mention the work of [14], which presents a stochastic EM algorithm for skewed  $\alpha$ -stable distributions.

The main contributions of this paper to the literature are as follows. First, we introduce new estimators for parameters of the  $\alpha$ -stable distribution, which mainly modify the well-known characteristic function estimator of this family by replacing the density function by a Gaussian kernel within the integral representation of the characteristic function. We also handle the estimation of the bandwidth parameter using the plug-in method proposed by [15] (for  $\alpha > 1$ ) and [16, 17] (for  $\alpha < 1$ ). Second, the ML estimator is computed numerically using score functions and the False position algorithm. Furthermore, this allows to improve the EM algorithm of [13]. The Bayesian approach of [18] is improved by combining Gibbs sampling and the Metropolis-Hastings algorithm. We consider two-component models motivated by applications, but it can be generalised to finite-component univariate  $\alpha$ -stable mixture models. The performance of the proposed estimators is compared with other candidates through a small-scale simulation study. All these methods have been used to estimate the reproduction number during the COVID-19 outbreak data in Mayotte. We also give other application of a dataset related to the distribution of N-acetyltransferase activity data in the blood of 245 unrelated individuals for a caffeine urinary metabolites involved in carcinogenic substances, which is available in [19] for two sub-populations study.

The paper is organised as follows. In section 2, we first recall some basic quantile methods originally developed by McCulloch [12] and introduce the kernel estimation method of the characteristic function and the estimation procedure for  $\alpha$ -stable distributions. We also consider a new framework for the ML estimation of  $\alpha$ -stable distributions based on the score function and the False position algorithm. Then, in section 3, we outline an adapted Bayesian and EM algorithm approaches to estimating the parameters of a two mixture of  $\alpha$ -stable distributions, which can be generalised to multiple mixtures. The performance of the above proposed estimators for estimating the parameters of  $\alpha$ -stable distributions is given in section 4. In section 5, we apply the proposed methods with the above mentioned COVID-19 outbreak data in Mayotte together with the distribution of the above already mentioned

N-acetyltransferase activity data for two sub-populations studies and a conclusion with future theoretical and practical developments, in particular new stochastic approximation methods and algorithms.

## 2 Parameter estimation of $\alpha$ -stable distributions

There are many approaches to estimate the parameters of  $\alpha$ -stable distributions. One of such approaches is the Fractional Lower Order Moments (FLOM) method, as described in [11]. This method considers the lower-order fractional moments of the data to estimate the parameters of the distribution. It is also less sensitive to outliers and does not require any iterative optimisation procedures. However, a limitation of the FLOM method is that it assumes that the data are independently and identically distributed, which may not be the case in some real-world applications. Furthermore, as shown in [11], it does not guarantee a good estimate of the skewness parameter  $\beta$ .

Another approach is the quantile method, improved by McCulloch [12], which is based on the relationship between the  $q$ -th quantile and the distribution parameters for any distribution. In the case of the  $\alpha$ -stable distribution, the  $q$ -th quantile depends on the scale and location parameters of the distribution as well as the  $\alpha$  and  $\beta$  parameters, it is also robust to outliers and can handle missing data, but it may not work well for distributions with slowly varying tails or non-smooth density functions. In the same context, a new estimation algorithm for the tail index was proposed in [20] by considering a quantile conditional variance ratio. Based on the explicit formula (1) of the characteristic function, one can establish an estimation method using regression (ECF) [2]. Although the lack of a closed-form expression for the probability density is a theoretical disadvantage, note that in practice they are computed numerically using an integral transformation. It is in this way that many authors perform the ML method for parameter estimation [2]. The aim of this section is to present the relevant methods that we consider in order to improve our main results.

First, we recall the quantile method originally developed by McCulloch, as described in [12], provided by the `libstable4u` package [21]. We will use such a method to initialise our algorithms but with the constraint  $\alpha \geq 0.6$  [12, 22]. The method has been simplified in recent years, and the version we use is based on the work presented in [2], which can be described as follows. Let  $x_p$  be the  $p^{\text{th}}$  quantile of the distribution  $X \sim S(\alpha, \beta, \gamma, \zeta; 0)$  and define the following quantities:

$$v_\alpha(\alpha, \beta, \gamma, \zeta) = \frac{x_{0.95} - x_{0.05}}{x_{0.75} - x_{0.25}}, \quad v_\beta(\alpha, \beta, \gamma, \zeta) = \frac{x_{0.05} + x_{0.95} - 2x_{0.5}}{x_{0.95} - x_{0.05}},$$

$$v_\gamma(\alpha, \beta, \gamma, \zeta) = x_{0.75} - x_{0.25}, \quad v_\zeta(\alpha, \beta, \gamma, \zeta) = -x_{0.5}.$$

In [12], McCulloch provides tables from which the values of the above parameters can be derived, considering a standard  $\alpha$ -stable distribution  $Z \sim S(\alpha, \beta, 1, 0; 0)$ . This is not restrictive, since the scaling property, as outlined in [1], allows to have  $x_p = \gamma z_p + \zeta$ , where  $z_p$  is the  $p^{\text{th}}$  quantile of  $Z$ . The value of the above parameters could then be deduced from these quantities using the following relationships [2]:

$$v_\alpha(\alpha, \beta, \gamma, \zeta) = v_\alpha(\alpha, \beta, 1, 0), \quad v_\beta(\alpha, \beta, \gamma, \zeta) = v_\beta(\alpha, \beta, 1, 0),$$

$$v_\gamma(\alpha, \beta, \gamma, \zeta) = \gamma v_\gamma(\alpha, \beta, 1, 0), \quad v_\zeta(\alpha, \beta, \gamma, \zeta) = \gamma v_\zeta(\alpha, \beta, 1, 0) - \zeta.$$

As we can see,  $v_\alpha$  and  $v_\beta$  are independent of the scale and location parameters, and the above relationships allow one to obtain a reliable estimate of the four parameters when the sample set is large, see [2]. In the following, the vector parameter  $\Theta_0 = (\alpha_0, \beta_0, \gamma_0, \zeta_0)$  will refer to the initial vector value parameter according to McCulloch's method. Now we will introduce the other methods.

## 2.1 Characteristic function-based estimation method

Let  $X \sim S(\alpha, \beta, \gamma, \zeta; 0)$  and let  $X_1, \dots, X_n$  be a sample of size  $n$  of  $X$  and  $F_n$  be the empirical cumulative distribution function:

$$\hat{\phi}_n^{(1)}(t) = \int_{\mathbb{R}} \exp(itx) dF_n(x) = \frac{1}{n} \sum_{j=1}^n \exp(itX_j).$$

Note that, given  $t \in \mathbb{R}$ ,  $\hat{\phi}_n^{(1)}(t)$  is a consistent estimator of  $\phi(t)$  for large values of  $n$  (by a simple application of the law of large numbers). We introduce an alternative estimation of the characteristic function using the kernel method. We are concerned with the Gaussian kernel for the case when  $\alpha > 1$ , expressed as  $K(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$  for all  $x \in \mathbb{R}$  and satisfying  $\int_{\mathbb{R}} K(z) dz = 1$ . This alternative estimate of the characteristic function is defined as follows:

$$\hat{\phi}_n^{(2)}(t) = \frac{1}{nh_n} \int_{\mathbb{R}} \exp(itx) \sum_{j=1}^n K\left(\frac{x - X_j}{h_n}\right) dx,$$

where the bandwidth satisfies  $\lim_{n \rightarrow \infty} h_n = 0$  in order to have the consistency. For practical purposes, and in the case of  $\alpha > 1$ , we choose the optimal bandwidth  $h_n$  by considering the plug-in method proposed by Sheather and Jones [15], since it is based on the mean square error. On the other hand (specifically for  $\alpha < 1$ ), we consider the second-generation plug-in method proposed by Slaoui in [16] (see also [17]). Set

$$\eta(\gamma t | \alpha; 0) = \frac{2}{\pi} t \log(\gamma |t|) \mathbb{I}_{\{\alpha=1\}} + \tan\left(\frac{\pi\alpha}{2}\right) \text{sign}(t) \gamma^{1-\alpha} (|t| - |t|^\alpha) \mathbb{I}_{\{\alpha \neq 1\}}. \quad (3)$$

We now define the following sample functions  $g_n(t)$  and  $h_n(t)$  to compute the estimated characteristic function at sample size  $n$ :

$$\begin{cases} g_n(t) = \Re(\hat{\phi}_n(t)), \\ h_n(t) = \Im(\hat{\phi}_n(t)), \end{cases}$$

where  $\hat{\phi}_n(t)$  could be  $\hat{\phi}_n^{(1)}(t)$  or  $\hat{\phi}_n^{(2)}(t)$ . We have:

$$\begin{cases} \sqrt{g_n(t)^2 + h_n(t)^2} = |\hat{\phi}_n(t)| = \exp(-\gamma^\alpha |t|^\alpha), \\ \arctan\left(\frac{h_n(t)}{g_n(t)}\right) = \arg(\hat{\phi}_n(t)) = -\gamma^\alpha \beta \eta(\gamma t | \alpha; 0) + \zeta t. \end{cases}$$

We also define the following quantities:

$$\begin{cases} y(t) &= \log(-\log(|\phi(t)|)) = \log(\gamma^\alpha) + \alpha \log(|t|), \\ z_k &= \arg(\hat{\phi}_n(t_k)), \\ B_k &= \hat{\gamma}^{\hat{\alpha}} \eta(\hat{\gamma} t_k | \hat{\alpha}; 0). \end{cases}$$

The relationship between  $y(t)$  and  $\log(|t|)$  can be modelled linearly with a slope of  $\alpha$  and an intercept of  $a = \log(\gamma^\alpha)$ . Set

$$y_k = \log(-\log(|\hat{\phi}_n(t_k)|)) = \log\left(-\log\left(\sqrt{g_n(t_k)^2 + h_n(t_k)^2}\right)\right),$$

where  $t_k$  is chosen on a grid of equally spaced points  $t_1, \dots, t_m$  using the sample data. We use the following weighted least squares method to minimise

$$\begin{aligned} S(a, \alpha) &= \sum_{k=1}^m W_k (y_k - a - \alpha \log(|t_k|))^2 \\ (\hat{a}, \hat{\alpha}) &= \arg \min_{(a, \alpha)} S(a, \alpha), \end{aligned}$$

where the weights  $W_k = \frac{1}{\sigma_k^2}$  and  $\sigma_k^2$  denotes the variance of the  $k$ th observation.

We get the following estimators:

$$\begin{aligned} \hat{\alpha}_m &= \left( \sum_{k=1}^m W_k \log(t_k) y_k - \frac{\sum_{k=1}^m W_k y_k}{\sum_{k=1}^m W_k} \times \sum_{k=1}^m W_k \log(t_k) \right) \\ &\quad \times \left( \sum_{k=1}^m W_k \log(t_k)^2 - \frac{\sum_{k=1}^m W_k \log(t_k)}{\sum_{k=1}^m W_k} \times \sum_{k=1}^m W_k \log(t_k) \right)^{-1}, \\ \hat{a}_m &= \frac{\sum_{k=1}^m W_j y_k - \hat{\alpha}_m \sum_{k=1}^m W_k \log(t_k)}{\sum_{k=1}^m W_j}, \\ \hat{\gamma}_m &= \delta_0 \exp\left(\frac{\hat{a}_m}{\hat{\alpha}_m}\right). \end{aligned}$$



Then from the following formula

$$z_k = \beta B_k + \zeta t_k,$$

again using the method of weighted least squares as described above, we obtain

$$\begin{aligned} \hat{\zeta}_m &= \gamma_0 \left( \sum_{k=1}^m W_k B_k z_k - \frac{\sum_{k=1}^m W_k t_k z_k}{\sum_{k=1}^m W_k t_k B_k} \times \sum_{k=1}^m W_k B_k^2 \right) \\ &\quad \times \left( \sum_{k=1}^m W_k B_k t_k - \frac{\sum_{k=1}^m W_k t_k^2}{\sum_{k=1}^m W_k t_k B_k} \times \sum_{k=1}^m W_k B_k^2 \right)^{-1} + \zeta_0, \\ \hat{\beta}_m &= \frac{\sum_{k=1}^m W_k t_k z_k - \hat{\zeta} \sum_{k=1}^m W_k t_k^2}{\sum_{k=1}^m W_k t_k B_k}. \end{aligned}$$

Since  $\hat{\phi}_n$  is consistent, the consistency of the estimated parameters  $\hat{\alpha}_m$ ,  $\hat{\beta}_m$ ,  $\hat{\gamma}_m$  could be obtained using the classical regression estimation method. A simulation study is carried out to evaluate the performance of the proposed estimation. The choice of  $m$  was taken from [23], which suggests choosing points  $t_k$  in the interval  $[0.1, 1]$ . We set the parameters  $\gamma_0$  and  $\zeta_0$  using the McCulloch method mentioned above [12].

## 2.2 Maximum Likelihood Approximation

In the following lines, we present a method for estimating  $\alpha$ -stable distributions within the framework of ML. Since the probability density function does not have a closed-form expression, the classical ML method is computationally difficult in this context, because the likelihood ratio does not exist explicitly. For this reason, we first use a numerical approximation of the density function of  $\alpha$ -stable distributions, which provides accurate estimates. We introduce the following method based on score functions and the False position algorithm, which is an intuitive way to estimate the underline parameters of  $\alpha$ -stable distributions. Suppose we are in parameterization 1 and  $\alpha > 1$ . Set

$$\begin{aligned} g_d(x|\alpha, \beta) &= \int_0^\infty \cos(xr + \beta\eta(r, \alpha; 1)) r^{d-1} \exp(-r^\alpha) dr \mathbb{I}_{\{0 < d < \infty\}} \\ &\quad + \int_0^\infty [\cos(xr + \beta\eta(r, \alpha; 1)) - 1] r^{d-1} \exp(-r^\alpha) dr \mathbb{I}_{\{-2 \min(1, \alpha) < d \leq 0\}}, \end{aligned}$$

$$\begin{aligned} \tilde{g}_d(x|\alpha, \beta) &= \int_0^\infty \sin(xr + \beta\eta(r, \alpha; 1)) r^{d-1} \exp(-r^\alpha) dr \mathbb{I}_{\{-\min(1, \alpha) < d < \infty\}} \\ &\quad + \int_0^\infty [\sin(xr + \beta\eta(r, \alpha; 1)) - xr] r^{d-1} \exp(-r^\alpha) dr \mathbb{I}_{\{\alpha > 1, -\alpha < d \leq -1\}}, \end{aligned}$$

$$h_d(x|\alpha, \beta) = \int_0^\infty \cos(xr + \beta\eta(r, \alpha; 1)) \log(r)r^{d-1} \exp(-r^\alpha) dr,$$

$$\tilde{h}_d(x|\alpha, \beta) = \int_0^\infty \sin(xr + \beta\eta(r, \alpha; 1)) \log(r)r^{d-1} \exp(-r^\alpha) dr,$$

for  $x \in \mathbb{R}$  and  $d \in \mathbb{N}$ . Note that  $\eta(r, \alpha; 1) = -\tan(\frac{\pi\alpha}{2}r^\alpha)$  should not be confused with the previous one in (3) from parameterization 0. Recall the following theorem from [2].

**Theorem 1** (Stable score function) *Let  $\alpha \neq 1$ . The univariate  $\alpha$ -stable density in the 1-parameterization is given by*

$$f(x|\alpha, \beta, \gamma, \zeta; 1) = \frac{1}{\pi\gamma} g_1 \left( \frac{x - \zeta}{\gamma} \middle| \alpha, \beta \right).$$

Then the score functions are given by

$$\begin{aligned} \frac{\partial f}{\partial \alpha}(x|\alpha, \beta, \gamma, \zeta; 1) &= \frac{1}{\pi\gamma} \left[ \frac{\pi\beta}{2 \cos(\frac{\pi\alpha}{2})^2} \tilde{g}_{1+\alpha} \left( \frac{x - \zeta}{\gamma} \middle| \alpha, \beta \right) \right. \\ &\quad \left. + \beta \tan\left(\frac{\pi\alpha}{2}\right) \tilde{h}_{1+\alpha} \left( \frac{x - \zeta}{\gamma} \middle| \alpha, \beta \right) - h_{1+\alpha} \left( \frac{x - \zeta}{\gamma} \middle| \alpha, \beta \right) \right], \\ \frac{\partial f}{\partial \beta}(x|\alpha, \beta, \gamma, \zeta; 1) &= \frac{\tan(\frac{\pi\alpha}{2})}{\pi\gamma} \tilde{g}_{1+\alpha} \left( \frac{x - \zeta}{\gamma} \middle| \alpha, \beta \right), \\ \frac{\partial f}{\partial \gamma}(x|\alpha, \beta, \gamma, \zeta; 1) &= -\frac{1}{\pi\gamma^2} g_1 \left( \frac{x - \zeta}{\gamma} \middle| \alpha, \beta \right) + \frac{x - \zeta}{\pi\gamma^3} \tilde{g}_2 \left( \frac{x - \zeta}{\gamma} \middle| \alpha, \beta \right), \\ \frac{\partial f}{\partial \zeta}(x|\alpha, \beta, \gamma, \zeta; 1) &= -\frac{1}{\pi\gamma^2} \tilde{g}_2 \left( \frac{x - \zeta}{\gamma} \middle| \alpha, \beta \right). \end{aligned}$$

Many equations, including most of the more complicated ones, can only be solved by iterative numerical approximation. There are many root-finding algorithms that can be used to obtain approximations to such a given root. One of the most common is Newton's method or the secant method, but it may fail to find a root under certain circumstances, and it can be computationally expensive since it requires a computation of the derivatives of the function. Other methods are needed and a general class of methods are the two-point bracket methods. The False position or Bisection algorithm is one of these. The convergence rate of the bisection method could possibly be improved by using a different solution estimate. The False position algorithm runs as the first iteration of the bisection algorithm and, essentially, the root is approximated by replacing the actual function by a line segment on the bracketing interval and then using the classic double False position formula on that line segment, see [24].

Obviously, this method requires a good range for each parameter. For this reason we consider the initial intervals with a given margin to the initial values  $\Theta_0 = (\alpha_0, \beta_0, \gamma_0, \zeta_0)$  based on the quantile method. The same procedure is

used for the other three parameters. In the following we will refer to this method as ML-Second. At this stage, however, one may be disappointed that there is no theoretical study of the existence of the root solutions of the above score functions under the False position method. This will be addressed as a perspective, but we will only look at the convergence numerically. Note that one can also use the `optim` command in `R` to directly maximise the log-likelihood function and obtain the estimated parameters. This is referred to as ML-first. The performance and consistency of the estimation is illustrated in a simulation study.

### 3 Parameter estimation for mixtures of $\alpha$ -stable distributions

Finite mixture models are becoming increasingly popular and play a crucial role in density estimation, and the mixture of  $\alpha$ -stable distributions is a popular tool for modelling skewed and impulsive data, making it applicable in various fields. A well-known method for mixture estimation is the EM algorithm. We will consider the ECF (2.1) and ML (2.2) methods for estimating the four parameters of these distributions, in order to select the most efficient of them for use in the EM algorithm. This algorithm is particularly useful in cases where the data are incomplete or partially observed. Much work has been done in this direction, including the papers cited by [13] and references therein. We should also mention the work of [25], [14], which presents a methodology for a stochastic EM algorithm applied to  $\alpha$ -stable distributions. In [26] the author established a formula that includes symmetrical and asymmetrical  $\alpha$ -stable distributions to estimate parameters using the EM algorithm, but the paper is still unpublished and we note that their algorithm does not converge for non-symmetric  $\alpha$ -stable distributions. Since the Bayesian approach is also a useful and efficient tool for parameter estimation in mixture models, we have proposed a modified framework that involves updating the posterior distribution until it converges to the stationary distribution by combining Gibbs sampling and the Metropolis-Hastings algorithm, as described in [18]. The novelty of this latest approach, compared to [18], is essentially the choice of the rejection zones in the Metropolis-Hastings algorithm, which, in addition to updating the parameters, can significantly affect the estimates. Mixture models in general could be used to understand the development an epidemic by estimating the generation time and the number of reproductions; see for example the work in [27], where the authors consider mixtures of Weibull, lognormal and other distributions to estimate the effective reproduction number during the COVID-19 outbreak on the island of Mayotte, France.

The mixture of  $\alpha$ -stable distributions is a popular tool for modelling skewed and impulsive data, making it applicable in various fields. When dealing with such a mixture method, one needs to initialise and select the components of the mixture. The K-means clustering [28] allows a better estimation of the components. Two common methods to infer parameters in

mixture models are: the EM algorithm and the Bayesian approach, [13, 18]. Bayesian estimation is a framework for formulating statistical inference problems. When predicting or estimating a random variable or process, the Bayesian philosophy is based on combining the evidence contained in the random variable with prior knowledge of the probability distribution of the random variable. The Bayesian estimation method uses prior data to estimate the value of the unknown parameters. This reduces the difference between the estimate and the true value of that parameter. In Bayesian modelling, the choice of priors then plays a crucial role in determining the posterior inference. The EM algorithm is a widely used computational method for estimating the parameters of statistical models with latent or missing variables. This algorithm is particularly useful in cases where the data are incomplete or partially observed. The EM algorithm works by iteratively alternating between the E-step, where we estimate the expected value of the unobserved or latent variables given the current parameter estimates, and the M-step, where we maximise the likelihood of the observed data based on the expected values obtained in the E-step. This alternating process continues until convergence is achieved, resulting in the optimal parameter estimates for the model. Let  $n$  be the number of observations and  $z_i$  the latent observations for all  $i = 1, \dots, n$ . Denote by  $\lambda_1 = \mathbb{P}(z_i = 1)$  the weight for the first component  $j = 1$  and  $\lambda_2 = 1 - \lambda_1$ . In this paper we assume that we are in a two-component mixture. Of course, this can be generalised to more than two components.

### 3.1 The proposed Expectation-Maximization algorithm

In this paper, we present an adapted Expectation-Maximisation algorithm by incorporating the above parameter estimation tools, namely the ML estimation method (by means of score functions and the False Position Algorithm) and the estimation by means of the Empirical or Kernel (ECF) function, both for the vector parameter  $\Theta$  and for updating the E-step in the EM algorithm. Such adjustments are important when selecting the appropriate parameter estimation method in the EM algorithm. For simplicity, we consider two sub-populations, say 1 and 2, and compute the values of the latent vector  $z$  as follows. A sample value  $x_i$  is assigned to population 1 with probability (the probability that  $z_i = 1$ , conditional on the observed value of  $x_i$ ), given by the key equation

$$\begin{aligned} p_i &= \mathbb{P}(z_i = 1|\Theta) \\ &= \frac{\lambda_1 f(x_i|\alpha_1, \beta_1, \gamma_1, \zeta_1; 0)}{\lambda_1 f(x_i|\alpha_1, \beta_1, \gamma_1, \zeta_1; 0) + (1 - \lambda_1) f(x_i|\alpha_2, \beta_2, \gamma_2, \zeta_2; 0)}. \end{aligned} \quad (4)$$

These are the actual weights assigned to observation  $i$  when calculating the expected log likelihood in the EM algorithm. Our proposed EM algorithm is given as follows. The subscript 't' indicates the t-th iteration of the EM algorithm to obtain the increment at each 'iteration'. We initialised  $\lambda_1^{(0)}$  using

the  $R$  software sample function noted `sample()` to draw a sample with attach specific probability to assign elements to subpopulation 1 or 2 according to the previous value in equation (4). We refer also to Remark 1 for other initialization options.

---

**Algorithm 1** EM algorithm for mixtures of  $\alpha$ -stable distributions
 

---

- 1: Initialization of the model, with selection of a specific tolerance  $\epsilon$ .
- 2: **repeat**
- 3:   **E-step:** We compute the  $n$  values of the vector  $z$ .
- 4:   **for**  $i = 1, \dots, n$  **do**
- 5:     A sample value  $x_i$  is assigned to the population 1 (and so we set  $z_i = 1$ ) with probability

$$P_i^{(t)} = \frac{\lambda_1^{(t)} f(x_i | \alpha_1^{(t)}, \beta_1^{(t)}, \gamma_1^{(t)}, \zeta_1^{(t)}; 0)}{\lambda^{(t)} f(x_i | \alpha_1^{(t)}, \beta_1^{(t)}, \gamma_1^{(t)}, \zeta_1^{(t)}; 0) + (1 - \lambda_1^{(t)}) f(x_i | \alpha_2^{(t)}, \beta_2^{(t)}, \gamma_2^{(t)}, \zeta_2^{(t)}; 0)}$$

- 6:   **end for**
  - 7:    $\lambda_1^{(t+1)} = \frac{1}{n} \{\#z_i = 1\}$ . So we use the expected log likelihood or ECF to get the new parameters for instance from component 1:  $\Theta_1^{(t+1)} = (\alpha_1^{(t+1)}, \beta_1^{(t+1)}, \gamma_1^{(t+1)}, \zeta_1^{(t+1)})$
  - 8:   **M-step:** We then independently maximise the new parameters for each case, for instance from the component 1 in the expected log-likelihood, which is denoted by  $Q^{(t+1)} = \sum_{i=1}^n \log (f(x_i | \alpha_1^{(t+1)}, \beta_1^{(t+1)}, \gamma_1^{(t+1)}, \zeta_1^{(t+1)}))$ , and the total log likelihood is the mean of the two cases.
  - 9: **until** convergence has been achieved :  $|Q^{(t+1)} - Q^{(t)}| < \epsilon$ .
- 

Note that, for a mixture model of two  $\alpha$ -stable distributions, it is identified in most cases by specifying  $\zeta_1 < \zeta_2$ . The EM algorithm thus fixes on one of the modes depending on its initialisation. In the Bayesian approach, one could simply include this in the priors for the two location parameters. The E-step is used to estimate and update the allocation parameter  $\lambda_1$  based on the values of the previous iteration, and once this is done, we can apply the usual estimation methods such as ML or ECF to the observations of each of the two components. In the M-step, we also maximise the conditional expectation for the well-chosen specific tolerance  $\epsilon$  until convergence. Incorporating the ECF method into the above EM algorithm has been shown to be a valid approach for mixtures of  $\alpha$ -stable distributions with suitable initialisation and specific tolerance  $\epsilon$ . Explicit and analytical studies for convergence by the ECF approach in the EM algorithm will be addressed as a perspective work. Let us now turn to another interesting method known for its flexibility, namely the Bayesian estimation method.

**Remark 1** *As an initialization of our proposed algorithm, we could implement our algorithm with a sufficient number of iterations and then check the convergence by*

plotting the values of the estimated parameters obtained at each iteration versus the iterations. A more rigorous option, consisted in using the initialization which maximize the likelihood.

### 3.2 The proposed Bayesian algorithm

As mentioned above, the Bayesian inference framework allows us to build a hierarchical model in which the unknown quantities are estimated via additional information (objective or not very informative prior or non-informative) or available data (informative priors) using Bayes' rule. The Bayesian inference has the advantage of providing credible intervals on the behaviour of the likelihood function, taking into account any given information on the parameters. The priors chosen for this model are as follows. We consider non-informative uniform priors for the exponent parameter  $\alpha$  and the skewness parameter  $\beta$  on their supports as in [29, 30]. Such choices are also discussed as being appropriate. An inverse gamma distribution with initial parameters (1, 1) is chosen for the dispersion  $\gamma$ , and a normal prior with parameters (0, 5) is chosen for the location parameter  $\zeta$ , and these priors are conjugate priors in Bayesian inference for the mean and variance [31]. Furthermore, in line with other work in the literature on mixing problems [32], the Beta distribution (the familiar 2-way special case of the Dirichlet distribution) is used as a prior for binomial proportions in Bayesian analysis [33]. We will consider the beta distribution  $Beta(1, 1)$  as a conjugate prior for the weights. Although it could be useful to use the information we know about the properties of the parameters of  $\alpha$ -stable distributions (e.g. they are bounded and as  $\alpha$  tends towards 2,  $\beta$  has less influence on asymmetry), the priors for  $\alpha, \beta, \gamma, \zeta$  were chosen inspired by [18] and assumed to be independent of each other to ensure that our model remains free of any unwanted biases. Speaking generally about other choices of priors, we hope to develop future research work over time by considering different choices of priors, such as Jeffrey's priors (well defined for the parameters of mixtures of distributions), which are a challenging task for  $\alpha$ -stable distributions since they are not available in closed form, especially when investigating the behaviour of the Fisher information matrix [3, 34, 35].

Note that computing the joint posterior distribution of the above parameters, given the data and priors, is often analytically intractable due to the lack of a closed-form expression for  $\alpha$ -stable densities. To overcome this problem, we will use Markov Chain Monte Carlo (MCMC) methods, more precisely a combination of Gibbs sampling and Metropolis-Hastings algorithms (see [18]), as illustrated in the following steps. The novelty of this latest approach, compared to [18], is essentially the choice of rejection zones in the Metropolis-Hastings algorithm, which can significantly affect the estimates in addition to the updating of the parameters. The standard deviation  $\sigma$  of the normal distribution used to select candidates in the Metropolis-Hastings procedure is 0.1. Any given choice of hyper-parameters for the priors will be updated in the

Metropolis-Hastings acceptance zone so that they are not significant for the speed of convergence with this approach.

### 3.2.1 The weight distribution

We assume that the prior distribution of the weights follows a  $Beta(\xi)$  distribution with initial parameters  $\xi = (1, 1)$ . Since  $\mathbb{P}(z_i = 1)$  is equal to  $\lambda_1$  for  $i = 1, \dots, N$ , where  $N$  is the number of observations, the full conditional distribution for  $\lambda = (\lambda_1, \lambda_2)$  is also a beta distribution, with updated parameters  $\xi_1 + n_1$  and  $\xi_2 + n_2$ , where for example  $n_i$  is the frequency of observations assigned to component  $i = 1, 2$ . Thus, the updated distribution for the weights is  $Beta(\lambda_1 + n_1, \lambda_2 + n_2)$ .

### 3.2.2 Updating the vector parameter $\Theta$ using MCMC

In this step we consider the Metropolis-Hastings sampling method. We generate a candidate parameter  $\Theta_j^{new} = (\alpha_j^{new}, \beta_j^{new}, \gamma_j^{new}, \zeta_j^{new})$ , for example  $j = 1$ , from a proposal distribution  $q(\cdot|\cdot)$ , and it is accepted with probability  $A_{\Theta_j^{new}}$ , defined by:

$$A_{\Theta_j^{new}} = \min \left( 1, \prod_{i=1, z_i=j}^N \frac{f(x_i | \alpha_j^{new}, \beta_j^{new}, \gamma_j^{new}, \zeta_j^{new}; 0)}{f(x_i | \alpha_j^{old}, \beta_j^{old}, \gamma_j^{old}, \zeta_j^{old}; 0)} \times \frac{p(\Theta_j^{new})q(\Theta_j^{old} | \Theta_j^{new})}{p(\Theta_j^{old})q(\Theta_j^{new} | \Theta_j^{old})} \right).$$

We also assume that the priors are independent. Then we get

$$p(\Theta_j) = p(\alpha_j)p(\beta_j)p(\gamma_j)p(\zeta_j).$$

In this paper we choose a normal distribution for  $q(\cdot|\cdot)$ . By symmetry we conclude that

$$q(\Theta_j^{new} | \Theta_j^{old}) = q(\Theta_j^{old} | \Theta_j^{new}).$$

Then  $A_{\Theta_j^{new}}$  become:

$$\min \left( 1, \prod_{i=1, z_i=j}^N \frac{f(x_i | \alpha_j^{new}, \beta_j^{new}, \delta_j^{new}, \omega_j^{new}; 0)}{f(x_i | \alpha_j^{old}, \beta_j^{old}, \delta_j^{old}, \omega_j^{old}; 0)} \times \frac{IG(\delta_j^{new} | \alpha_0, \beta_0)N(\omega_j^{new} | \epsilon, k)}{IG(\delta_j^{old} | \alpha_0, \beta_0)N(\omega_j^{old} | \epsilon, k)} \right). \quad (5)$$

Now we sample a uniform variable  $u$  in  $[0, 1]$ . If  $A_{\Theta_j^{new}} > u$ , we accept the new candidate variables, otherwise we keep those from the previous iteration. The fact that we consider a single rejection zone associated with the vector parameter  $\Theta$  is possible because the priors are assumed to be independent. Thus, the Markov chain  $\Theta_n = (\alpha_n, \beta_n, \gamma_n, \zeta_n)$ , where  $n$  is the iteration index, stationary, unlike [18], where the authors consider a multiple Markov chain for each parameter without taking advantage of independence.

### 3.2.3 Updating the allocation parameter

At each iteration, it is necessary to predict which subpopulation each observation belongs to. We do this by computing the conditional probability, for example for  $j = 1$  as in (4), which is the probability that the observation  $x_i$  belongs to the component  $j = 1$ . Note that this method requires ordered steps to converge to the correct distribution, similar to the approach described in [18], where reversible Markov chain Monte Carlo was used to determine the number of components in the mixture model. However, unlike the approach in [18], we consider equation (5) as the rejection zone in the Metropolis-Hastings step for all parameters, which seems to be numerically more accurate (considering the bias generated) than considering rejection zones for each parameter separately. We therefore summarise our adapted method in the following algorithm:

---

**Algorithm 2** Bayesian algorithm for mixtures of  $\alpha$ -stable distributions

---

**Require:** Initialization of weight parameters.

**Require:** Number of iterations  $N$  and burn-in  $M$ .

- 1: **for**  $t = 1, \dots, N$  **do**
  - 2:   Obtain weights  $\lambda = (\lambda_1, \lambda_2)$  by drawing samples from a symmetric beta distribution  $\lambda \sim \text{Beta}(\xi_1 + n_1, \xi_2 + n_2)$  where  $n_1$  is the frequency of observations assigned to the first component and  $n_2$  to the second.
  - 3:   Update the parameters of the proposal distribution  $q(\cdot|\cdot) = N(\cdot|\theta, \sigma)$ , setting  $\theta$  to the value of the previous iteration and choosing a small value for  $\sigma$  (the standard deviation).
  - 4:   Sample new candidates  $\Theta_j^{\text{new}} = (\alpha_j^{\text{new}}, \beta_j^{\text{new}}, \gamma_j^{\text{new}}, \zeta_j^{\text{new}})$  from the proposal distribution  $q(\cdot|\cdot) = N(\cdot|\theta, \sigma)$  for each component.
  - 5:   Accept  $\Theta_j^{\text{new}}$  according to equation (5) and set  $\Theta_j^t = \Theta_j^{\text{new}}$ , otherwise set  $\Theta_j^t = \Theta_j^{t-1}$ .
  - 6:   **for** each observation  $x_i$  **do**
  - 7:     Obtain the allocation variable  $z_i$  using equation (4).
  - 8:   **end for**
  - 9: **end for**
  - 10: Compute the mean parameters:  $\Theta_j = \frac{1}{N-M} \sum_{k=M}^N \Theta_j^{(k)}$ .
- 

In the Bayesian algorithm above, we compute the mean of the posterior rather than its median and confidence intervals. In fact, the parameter vector of the mixture of  $\alpha$ -stable distributions in the Bayesian method is a convergent Markov chain. Taking the median as an approximation for the parameter vector leads to numerical problems and there's no guarantee that the chain will converge. To numerically overcome such problems, we used Monte Carlo methods to estimate the parameter vector, which proved its effectiveness in the simulation part with well-selected burn-in period  $M$ .



Now that we have established the main methodology of interest in this paper, the next section will focus on assessing the effectiveness of each approach using simulated data and then applying it to real data.

## 4 Simulations

Our primary objective in this section is to evaluate the effectiveness of these approaches proposed in Sections 3.1 and 3.2 by using simulated data. We first discuss the performance of our proposed estimator based on the ECF and ML methods for estimating the parameters of single  $\alpha$ -stable distributions. Secondly, we consider the case of mixture of  $\alpha$ -stable distributions with different parameters. Bold values indicate smallest values closed to the true value in their respective rows when accessing the simulation performance. To have more precisions about the simulation performance we use the Mean Square Error (MSE) over the entire set of vector parameters  $\Theta$ :

$$MSE = \frac{1}{\dim(\Theta)} \sum_{i=1}^{\dim(\Theta)} (\hat{\theta}_i - \theta_i)^2.$$

### 4.1 The case of single $\alpha$ -stable distributions

We evaluate the performance through a simulation study of these methods and assess the effect of the observation size  $n$  with varying parameter values. The results of these simulations are presented through Table 1 to ???. We consider the Gaussian kernel in the ECF method (denoted by ECF-Kernel) and the ECF method with empirical function is denoted by ECF-Empirical, as presented above. In addition, we used the ML method (First and Second). It is worth noting that when using the ML-Second method, one observes that numerical convergence holds for the case  $\alpha > 1$  as mentioned in the methodology. For the case  $\alpha < 1$ , we have not been able to show efficient results because we had problems with the convergence of the special functions  $g_d$ ,  $h_d$ ,  $\tilde{g}_d$  and  $\tilde{h}_d$ . This will be evaluated in more detail in further work. According to Table 1 through Table 8, we find that the estimation methods demonstrated satisfactory performance.

**Table 1** Comparison of parameter estimation methods - configuration 1. **Bold values** indicate smallest values closed to the true value in their respective rows.

Parameter	True value	n	ECF-Kernel	ECF-Empirical	ML-First	ML-Second
$\alpha$	1.6	500	1.6356	<b>1.6023</b>	1.6078	1.6151
		750	<b>1.6191</b>	1.5484	1.5697	1.7641
		1000	<b>1.6075</b>	1.5519	1.5443	1.6894
$\beta$	-0.8	500	-0.7541	<b>-0.7566</b>	-0.6692	-0.3714
		750	-0.8375	-0.7668	-0.7389	<b>-0.8081</b>
		1000	-0.7696	-0.7337	<b>-0.7781</b>	-0.7756
$\gamma$	5	500	<b>5.0124</b>	4.8583	4.8791	5.0498
		750	5.2009	<b>4.9608</b>	4.9314	5.4934
		1000	5.1666	<b>4.9888</b>	4.9418	5.2904
$\zeta$	12	500	11.8914	<b>11.9563</b>	11.8727	12.8558
		750	<b>12.2187</b>	12.2930	12.2241	12.8333
		1000	<b>12.1193</b>	12.1905	12.2968	12.5753

**Table 2** Mean square error - configuration 1. **Bold values** indicate smallest values of MSE in their respective rows

n	ECF-Kernel	ECF-Empirical	ML-First	ML-Second
500	<b>0.0038</b>	0.0059	0.0119	0.2296
750	0.0224	0.0227	<b>0.0148</b>	0.2412
1000	<b>0.0107</b>	0.0108	0.0237	0.1059

**Table 3** Comparison of parameter estimation methods - configuration 2.

Parameter	True value	n	ECF-Kernel	ECF-Empirical	ML-First	ML-Second
$\alpha$	1.4	500	<b>1.3848</b>	1.3243	1.3588	1.3605
		750	<b>1.3973</b>	1.3613	1.3649	1.4420
		1000	1.3329	1.2985	1.3210	<b>1.3924</b>
$\beta$	0.5	500	0.4304	0.4464	<b>0.5150</b>	0.5800
		750	<b>0.5173</b>	0.5651	0.5523	0.6135
		1000	0.5381	0.5542	<b>0.5280</b>	0.6106
$\gamma$	2	500	<b>1.9819</b>	1.8854	1.9495	1.9278
		750	2.1368	<b>2.0573</b>	2.0668	2.2048
		1000	2.0910	<b>2.0162</b>	2.0397	2.2395
$\zeta$	-10	500	-9.9746	-10.0041	<b>-10.0032</b>	-8.4292
		750	-9.9876	-10.0300	<b>-10.0021</b>	-9.2277
		1000	-10.0807	-10.1048	<b>-10.0705</b>	-10.8141

**Table 4** Mean square error - configuration 2.

n	ECF-Kernel	ECF-Empirical	ML-First	ML-Second
500	0.0015	0.0054	<b>0.0011</b>	0.6201
750	0.0047	0.0024	<b>0.0021</b>	0.1632
1000	0.0051	0.0061	<b>0.0033</b>	0.1831

**Table 7** Comparison of parameter estimation methods - configuration 4.

Parameter	True value	n	ECF-Kernel	ECF-Empirical	ML-First
$\alpha$	0.6	500	0.7295	0.6118	<b>0.5932</b>
		750	0.7136	<b>0.5949</b>	0.5852
		1000	0.6826	0.6447	<b>0.5801</b>
$\beta$	-0.5	500	-0.3862	<b>-0.5280</b>	-0.5449
		750	-0.3669	-0.4413	<b>-0.4831</b>
		1000	-0.4074	<b>-0.4941</b>	-0.4436
$\gamma$	4	500	4.7518	<b>4.0385</b>	4.0950
		750	4.3565	3.7984	<b>3.8458</b>
		1000	4.4246	<b>4.2515</b>	3.7425
$\zeta$	5	500	4.1715	4.4790	<b>4.8560</b>
		750	4.5660	4.7277	<b>5.1534</b>
		1000	4.7741	<b>4.9808</b>	5.1215

**Table 5** Comparison of parameter estimation methods - configuration 3.

Parameter	True value	n	ECF-Kernel	ECF-Empirical	ML-First
$\alpha$	0.8	500	0.9831	0.8523	<b>0.7717</b>
		750	0.8601	<b>0.7891</b>	0.7730
		1000	0.8614	0.7502	<b>0.7510</b>
$\beta$	0.8	500	<b>0.7755</b>	0.7249	0.8436
		750	0.8903	0.9398	<b>0.8166</b>
		1000	0.6330	0.7373	<b>0.8056</b>
$\gamma$	3	500	3.3389	3.1960	<b>3.0887</b>
		750	3.2231	<b>3.0899</b>	3.0994
		1000	3.0905	2.8824	<b>2.9579</b>
$\zeta$	-12	500	-11.5866	-11.6107	<b>-11.9937</b>
		750	<b>-12.0331</b>	-12.1348	-12.0439
		1000	<b>-11.9651</b>	-12.0894	-12.1991

**Table 6** Mean square error - configuration 3.

n	ECF-Kernel	ECF-Empirical	ML-First
500	0.0799	0.0495	<b>0.0026</b>
750	0.0156	0.0114	<b>0.0032</b>
1000	0.0102	<b>0.0070</b>	0.0109

**Table 8** Mean square error - configuration 4.

n	ECF-Kernel	ECF-Empirical	ML-First
500	0.3203	0.0684	<b>0.0079</b>
750	0.0865	0.0295	<b>0.0119</b>
1000	0.0616	0.0164	<b>0.0211</b>

## 4.2 The case of mixture of $\alpha$ -stable distributions

We consider here, the framework of mixture estimation of two  $\alpha$ -stable distributions with different parameters. The aim is to evaluate the accuracy and efficiency of the methods described above, namely our adapted EM algorithm (including the use of ML-First, ML-Second, ECF-Kernel, ECF-Empirical) to update the parameters in the M-step, and the Bayesian method. Let  $\Sigma_1 = (\alpha_1, \beta_1, \gamma_1, \zeta_1)$  and  $\Sigma_2 = (\alpha_2, \beta_2, \gamma_2, \zeta_2)$ ; the density of the mixture model is given by:

$$f(x, \lambda, \Sigma_1, \Sigma_2; 0) = \lambda \times f(x, \Sigma_1; 0) + (1 - \lambda) \times f(x, \Sigma_2; 0).$$

After applying the four methods, we obtain the following table of estimated parameters:

According to Table 9 through Table ??, we find again that, all the estimation methods demonstrated satisfactory performance. The corresponding graphs to visualise the plots are presented through Figure 1 to 11.

**Table 9** Comparison of methods for estimating mixture model-configuration 0. **Bold values indicate smallest values in their respective rows.**

Parameter	True value	Bayesian	EM-ECF-Kernel	EM-ECF-Empirical	EM-ML
$\alpha_1$	1.2	<b>1.2032</b>	1.2099	1.2886	1.1689
$\beta_1$	0.5	<b>0.5015</b>	0.2977	0.4095	0.4906
$\gamma_1$	1	<b>0.9988</b>	1.0338	1.0253	0.9638
$\zeta_1$	-4.25	<b>-4.1939</b>	-4.1146	-4.1474	-4.1906
$\lambda_1$	0.6	0.6037	0.622	<b>0.599</b>	<b>0.601</b>
$\alpha_2$	1.2	<b>1.1812</b>	1.4926	1.1579	1.1434
$\beta_2$	-0.5	-0.6354	-0.6280	-0.3839	<b>-0.5497</b>
$\gamma_2$	0.5	<b>0.5062</b>	0.5155	0.5103	0.4887
$\zeta_2$	0.3	<b>0.3043</b>	0.2694	0.2808	0.3157

**Table 10** Comparison of methods for estimating mixture model-configuration 1. **Bold values indicate smallest values in their respective rows.**

Parameter	True value	Bayesian	EM-ECF-Kernel	EM-ECF-Empirical	EM-ML
$\alpha_1$	1.2	1.1206	1.1897	1.1804	1.1170
$\beta_1$	0.5	<b>0.4741</b>	0.3107	0.4041	0.5095
$\gamma_1$	1	<b>0.9805</b>	0.9959	1.0263	1.0309
$\zeta_1$	-2.5	<b>-2.4334</b>	-2.3928	-2.3867	-2.4098
$\lambda_1$	0.6	0.6166	<b>0.602</b>	0.617	0.632
$\alpha_2$	1.7	<b>1.6958</b>	1.7426	1.7412	1.7488
$\beta_2$	0.5	<b>0.4871</b>	0.9308	0.8949	0.9999
$\gamma_2$	0.8	0.7604	<b>0.8348</b>	0.7400	0.7011
$\zeta_2$	3	2.9495	2.9409	2.9855	<b>2.9730</b>

**Table 11** Comparison of methods for estimating mixture model-configuration 2. **Bold values indicate smallest values in their respective rows.**

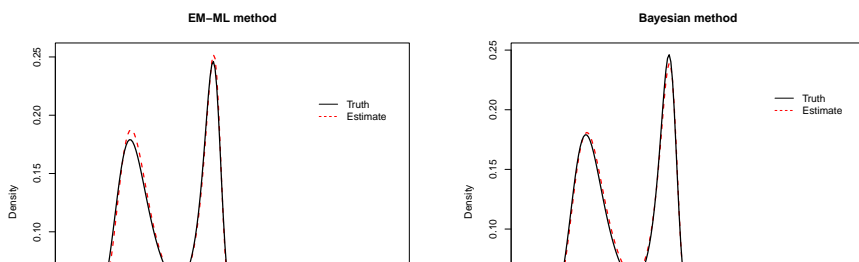
Parameter	True value	Bayesian	EM-ECF-Kernel	EM-ECF-Empirical	EM-ML
$\alpha_1$	1.4	<b>1.4115</b>	1.4655	1.4116	1.4969
$\beta_1$	-0.5	<b>-0.5134</b>	-0.5944	-0.6479	-0.9999
$\gamma_1$	0.8	<b>0.8259</b>	0.8302	0.7766	0.7787
$\zeta_1$	1.5	<b>1.5957</b>	1.6116	1.5991	1.6686
$\lambda_1$	0.5	0.6072	0.6190	0.6080	<b>0.6060</b>
$\alpha_2$	1.8	1.7398	1.8292	<b>1.8203</b>	1.6979
$\beta_2$	0.5	<b>0.3467</b>	1	1	0.9999
$\gamma_2$	1.2	<b>1.2296</b>	1.1782	1.1437	1.0805
$\zeta_2$	5.3	5.4898	5.4519	<b>5.3541</b>	5.4004

**Table 12** Mean square error for estimating mixture model-configuration 0.

Bayesian	EM-ECF-Kernel	EM-ECF-Empirical	EM-ML
0.0054	0.0410	0.0107	<b>0.0029</b>

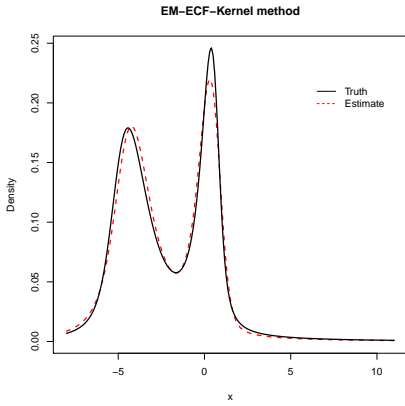
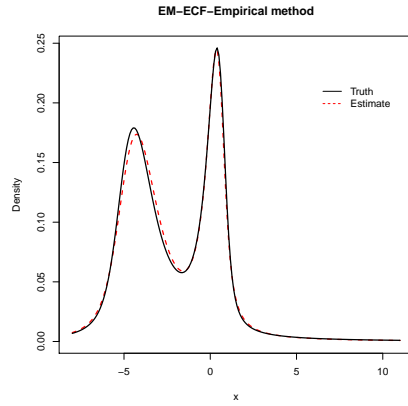
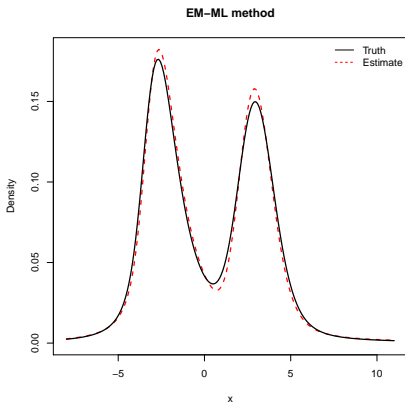
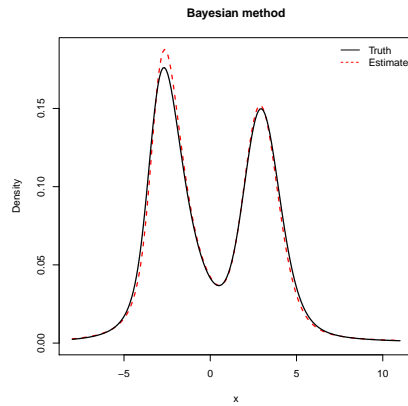
**Table 13** Mean square error for estimating mixture model-configuration 1.

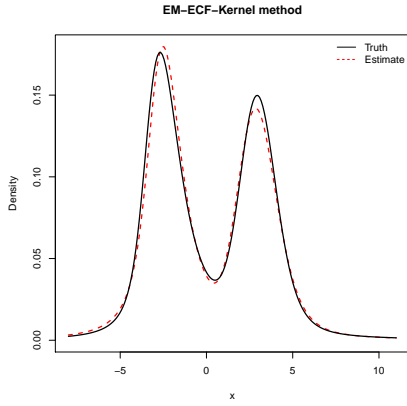
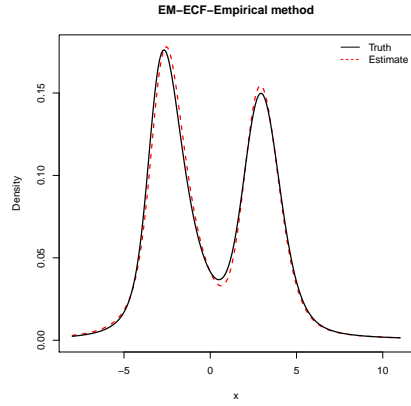
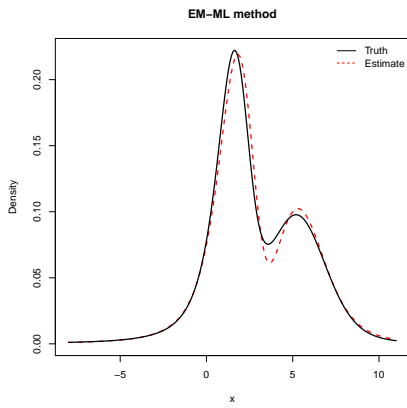
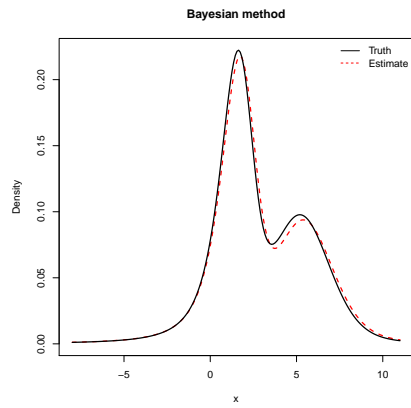
Bayesian	EM-ECF-Kernel	EM-ECF-Empirical	EM-ML
<b>0.0018</b>	0.0266	0.0205	0.0310

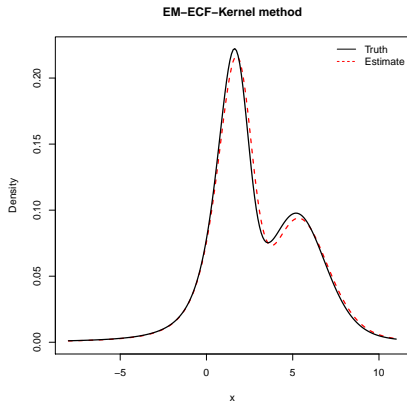
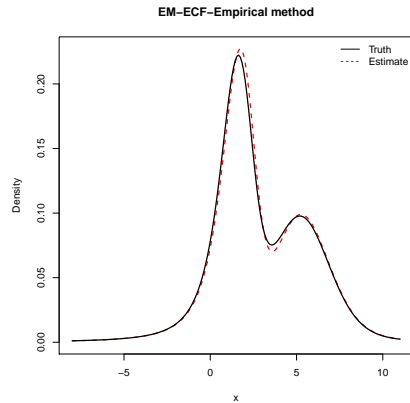


**Table 14** Mean square error for estimating mixture model-configuration 2.

Bayesian	EM-ECF-Kernel	EM-ECF-Empirical	EM-ML
<b>0.0095</b>	0.0350	0.0333	0.0648

**Fig. 3** EM-ECF-Kernel-configuration 0**Fig. 4** EM-ECF-Empirical-configuration 0**Fig. 5** EM-ML-configuration 1**Fig. 6** Bayesian-configuration 1

**Fig. 7** EM-ECF-Kernel-configuration 1**Fig. 8** EM-ECF-Empirical-configuration 1**Fig. 9** EM-ML-configuration 2**Fig. 10** Bayesian-configuration 2

**Fig. 11** EM-ECF-Kernel-configuration 2**Fig. 12** EM-ECF-Empirical-configuration 2

The above simulation performance shows numerically accurate parameter estimation for the mixture of  $\alpha$ -stable distributions with a given two components, which can be extended to more than two components.

## 5 Applications and conclusions

Let us now propose an application of the methods described above to real data, namely the serial interval (SI) in epidemiology, which is the time interval between the onset of symptoms in the infector and the onset of symptoms in the infectee. The generation time interval is then derived as the time interval from the infection of the infector to infection of the infectee. Thus, it is the time lag between infection in a primary case and a secondary case; and should be obtained from the time lag between all infectee/infector pairs [36]. As it cannot be observed directly, it is often replaced by the SI. Estimating the SI generation time and effective reproductive number [27] is an important task in understanding and developing an epidemic. In the previous paper [27] we only consider non-negative SI from the serial interval dataset of [37]. However, this dataset also contains negative serial intervals, because a suspected infector may show symptoms (infection) only after the infected person does. We also consider a data set related to the distribution of caffeine as a probe drug to determine the genetic status of two subpopulations of fast and slow acetylators. Acetylator status was determined from the urinary metabolic ratio for an enzymatic activity in the blood involved in the metabolism of carcinogenic substances, available in [38] for 245 unrelated individuals. Mixture models have been applied to the distribution of the above Acetylator status dataset as it contains two subpopulations of slow and fast metabolizers within the population. In the case of the above SI dataset, although the histogram in Figure 17 shows a mode in the last bin around 18, the K-means method confirms the choice of a two-component mixture. We perform a goodness-of-fit test to

determine whether or not our sample data fits a normal distribution. As not all points lie approximately on the reference line in Figure 18, we can assume that the distribution of our data sets is not normally distributed (see also Table 19 for the Shapiro-Wilk or Anderson-Darling or Kolmogorov-Smirnov tests). In addition, we use skewness and kurtosis, which are two important measures in statistics; skewness refers to the lack of symmetry and kurtosis is a measure of whether or not a distribution is heavy-tailed. To calculate the skewness and kurtosis of our data sets, we use the `skewness()` and `kurtosis()` functions from the `moments` library in R software. This allows us to see that our dataset distributions are skewed and not symmetric (see Table 20). Using the Kurtosis, we see that the distribution is leptokurtic showing heavy tails. We notice that the distributions of our data sets are skewed, not symmetric, and have heavy tails. Therefore, we proposed to fit the above real data distributions with  $\alpha$ -stable (mixture) distributions. Note, however, that some statistical goodness-of-fit tests designed for  $\alpha$ -stable distributions are available in the literature and require large sample sizes [39–41]. In particular, the recent work of [42] proposes a novel goodness-of-fit method based on quantile (trimmed) conditional variances.

Tables and comparisons of AIC and BIC criteria are provided to enhance understanding of the estimated parameters. The COVID-19 pandemic has caused considerable damage worldwide, disrupting productivity and causing widespread panic. In the French region of Mayotte, the regional health authority has made significant efforts to collect and monitor the spread of COVID-19, as documented in [27], by estimating the time-varying reproduction number, which is a non-pharmaceutical monitoring tool. The reported temporal daily cases of COVID-19 from 13 March 2020 to 11 January 2022 are shown in Figure 16.

We found that the EM-ML method performed well in our analysis due to the small size of the data set, the peaked distribution and, the fact that the assignment vector changes at each iteration, which can affect the estimation of model parameters. This is also consistent with the work of [13]. To validate our decision, we calculated the AIC and BIC criteria for all four methods that led to this optimal choice. Although the AIC and BIC values were relatively close, we visually confirmed this choice in Figure 4.2. The tables associated with these methods are presented in Tables 15 and 16.

The estimated curves obtained by combining each method are shown in Figure 13. It can be seen that the mixture of two  $\alpha$ -stable distributions using the EM-ML method provides the best fit of the serial interval.

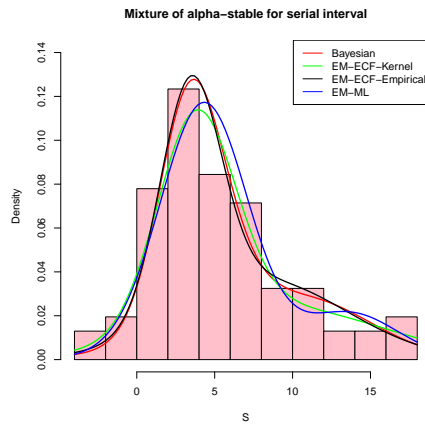


**Table 15** Estimated parameters for the serial interval distribution above

Parameter	Bayesian	EM-ECF-Kernel	EM-ECF-Empirical	EM-ML
$\alpha_1$	1.7140	2	2	1.9614
$\beta_1$	0.4010	1	1	-0.8793
$\gamma_1$	1.6734	1.7524	1.3736	2.0038
$\zeta_1$	3.6323	3.8096	3.4368	4.2990
$\mu_1$	3.9558	3.8096	3.4368	4.2990
$\lambda$	0.7111	0.5974	0.5064	0.8311
$\mu = \lambda\mu_1 + (1 - \lambda)\mu_2$	5.7253	5.9183	5.7112	5.8660
$\alpha_2$	1.7625	2	2	2
$\beta_2$	-0.6113	-1	-1	0
$\gamma_2$	3.2759	4.7511	3.8474	2.2267
$\zeta_2$	10.8648	9.0474	8.0447	13.5769
$\mu_2$	10.08101	9.0474	8.0447	13.5769

**Table 16** Comparison table between the selection criteria for the serial interval distribution above

Methods	EM-ML	EM-ECF-Kernel	EM-ECF-Empirical	Bayesian
AIC	<b>455.2616</b>	457.9891	455.935	456.9538
BIC	<b>476.3558</b>	479.0833	477.0293	478.048

**Fig. 13** Mixture estimation of the serial interval distribution

## 5.1 The effective reproduction number

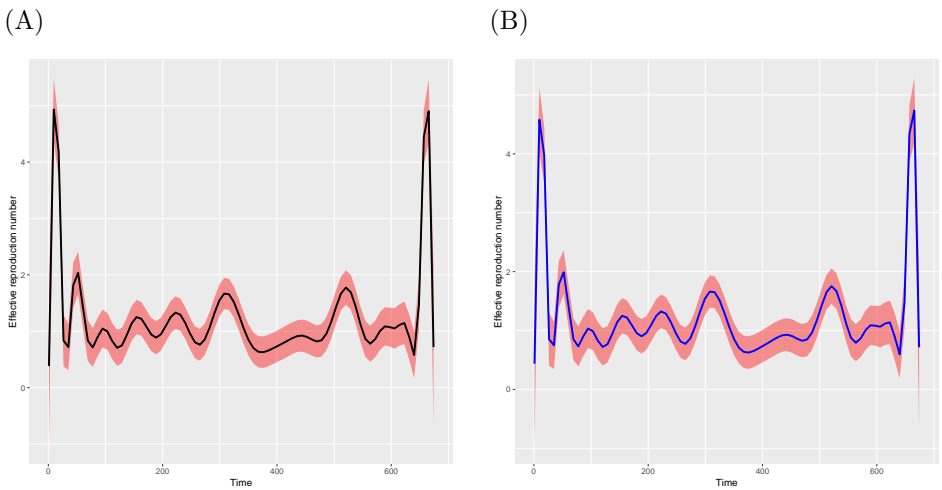
The basic reproduction number  $R_0$  at the start of an epidemic and the time varying (effective) reproduction numbers during an outbreak are important tools. Historically, it has been defined as the average number of new infections generated by an individual during a period of infectivity (see [43]). There are several methods to calculate this parameter. We consider a non-parametric approach [44] based on the generation time function associated with the serial interval distribution (see [45, 46]). In the previous paper [27] we have

only considered non-negative SI subset data from the [37] dataset and looked for SI estimation models such as Gamma, Lognormal and Weibull. Here we consider the entire dataset, including negative SI values, and deal with the  $\alpha$ -stable mixture modelling framework. This is not a new dataset, but an original dataset from [37].

Let  $p$  be the probability distribution of the transmissibility of an infectious individual at age of infection  $\tau$ , assuming that the entire population is susceptible. Let  $\Gamma(t)$  be the number of new infections during the time interval  $]t; t + dt[$ . For discrete time  $t \neq 0$ , we have the following non-parametric formula for the effective reproduction number:

$$R_0(t) = \frac{\Gamma(t)}{\sum_{\tau \leq t} p(\tau) \Gamma(t - \tau)}. \quad (6)$$

Other improved methods of estimating the effective reproductive number exist in the literature, see [47, 48] and references therein. In this work, using the epidemic incidence curve in Mayotte (between 13 March 2020 and 11 January 2022), we derive a generation time distribution to estimate the effective reproduction number  $R_0(t)$  using the non-parametric formula in equation (6) and the best fit of the SI in a  $\alpha$ -stable mixture modelling framework. The following plots in Figure 14 show the evolution of the time-varying reproduction number  $R_0(t)$ ; we smooth the curve using estimated values.



**Fig. 14** Evolution of the effective reproductive number in Mayotte from 13 March 2020 to 11 January 2022 with a mixture of  $\alpha$ -stable distributions with a serial interval estimated by EM-ML (A) and estimated by EM-ECF-Empirical (B).

If we look at the graphs of Figure 14, we see a complete similarity between them, and for both we can see that the effective reproduction number starts

at a value of around 2.5, indicating that each infected individual infects, on average, 2–3 other people. Over time, we see a sharp increase in the reproduction number, peaking at around 4.5, indicating that the disease is spreading rapidly. This could be due to a number of factors, such as increased travel, relaxed social distancing measures, or a new variant of the disease that is more transmissible. After the peak, we see a decline in the number of reproductive cases, indicating that the disease is spreading more slowly. This could be due to interventions such as increased vaccination rates, stricter social distancing measures, or natural immunity acquired by those who have recovered from the disease. The reproduction number will eventually fall below 1, indicating that the disease is no longer spreading and can be considered to be under control. Eventually, however, the reproduction number will suddenly increase due to the effect of the new omicron variant.

## 5.2 The distribution of N-acetyltransferase activity data

Here we consider a dataset relating to the distribution of N-acetyltransferase activity in the blood of  $n = 245$  unrelated individuals for a caffeine urine metabolite test for an enzyme involved in the metabolism of carcinogens available in Bethtel et al. [19]. This data set has been used in the past to test mixture models as it contains two subpopulations of slow and fast metabolisers within the population.

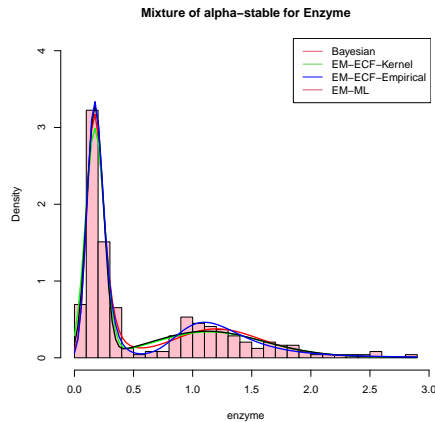
**Table 17** Table of estimated parameters for the above distribution of N-acetyltransferase activity data

Parameter	Bayesian	EM-ECF-Kernel	EM-ECF-Empirical	EM-ML
$\alpha_1$	1.5936	1.9857	2	1.6939
$\beta_1$	0.8501	1	1	0.9999
$\gamma_1$	0.0549	0.0565	0.0499	0.0525
$\zeta_1$	0.1756	0.1726	0.1741	0.1752
$\mu_1$	0.2102	0.1738	0.1741	0.2025
$\lambda$	0.6188	0.5918	0.5755	0.6244
$\mu = \lambda\mu_1 + (1 - \lambda)\mu_2$	0.6273	0.5939	0.6059	0.6801
$\alpha_2$	1.7175	1.8921	1.8864	1.3883
$\beta_2$	0.6134	1	1	0.9964
$\gamma_2$	0.2885	0.3409	0.3458	0.2273
$\zeta_2$	1.2204	1.1447	1.1292	1.1502
$\mu_2$	1.3045	1.2030	1.1915	1.4742

**Table 18** Table comparing the selection criteria for the above distribution of N-acetyltransferase activity data

Methods	EM-ML	EM-ECF-Kernel	EM-ECF-Empirical	Bayesian
AIC	<b>109.2265</b>	129.2364	129.3528	116.7802
BIC	<b>140.7378</b>	160.7478	160.8641	148.2916

We observe that almost all the estimated parameters are similar, except for the stability index  $\alpha$ , see Table 17. The EM-ML model is the most effective, as shown in Table 18, with the lowest AIC and BIC values. In addition, the corresponding figure for the four methods is shown in Figure 15.



**Fig. 15** Mixture methods applied to the distribution of N-acetyltransferase activity data in the blood of 245 unrelated individuals in Bethtel et al. [19].

The EM-ML method provides a better fit to the distribution of N-acetyltransferase activity in the blood data than the other approaches. This is because the other methods, particularly EM-ECF, are sensitive to changes in the assignment vector at each iteration. Furthermore, the Bayesian approach still has the advantage of using any prior information on the parameters, which explains its efficiency on simulated data. However, it also requires certain conditions, such as the length of the observations, to ensure accurate results.

### 5.3 Conclusion and perspectives

In this paper, we consider the parameter estimation of univariate  $\alpha$ -stable distributions and their mixture. We introduce some new techniques, such as the Gaussian kernel estimator in the characteristic function for the case  $\alpha \geq 1$ , which has shown more efficient performance than the empirical characteristic function in the simulation study. We also perform another estimation procedure in the ML framework based on the [False position](#) algorithm method to find the root of the log-likelihood through the score functions established in [2].

In the case of estimating the mixture of  $\alpha$ -stable distributions, although we have limited our analysis to two components, the proposed methods can be generalised to multiple components. The EM algorithm was adapted to estimate the parameters of each subpopulation to convergence by combining the ECF and ML methods in the M-step. Integrating the ECF into the M stage has proved to be a valuable practical approach.

The Bayesian method, which is more flexible but requires many steps to perform the estimation, has also been adapted to the parameter estimation of a given mixture model of  $\alpha$ -stable distributions. Finally, we consider two types of applications of our estimation methods on real data, namely the estimation of the reproduction number of the COVID-19 in Mayotte and the Acetylator status dataset.

In terms of future research, an immediate practical study is the implementation of an R package for parameter estimation of a mixture of  $\alpha$ -stable distributions as well as a focus on statistical goodness-of-fit tests designed for mixture of  $\alpha$ -stable distributions. We also plan to implement a more efficient approach to overcome the long burn-in period in the MCMC method using importance sampling [49]. Again, a future research direction would be to extend our study to the case of non symmetric  $\alpha$ -stable distributions using Gibbs sampling [50].

We also aim to study some identity representations for  $\alpha$ -stable distributions that incorporate a Weibull location scale mixture model, as stated in [26] in the symmetric case. In this way, our goal is to develop a stochastic algorithm that can account for the additional computational complexity [51] of the Weibull distribution.

## Supplementary information

Additional supporting information on the original papers presented in the study and the R codes including datasets are available from the corresponding author on reasonable request.

## Acknowledgements

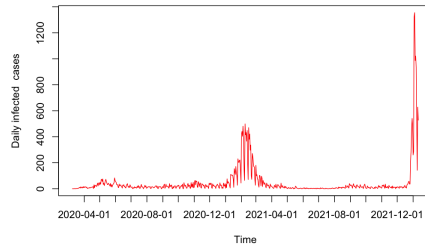
The authors thank the reviewers whose comments improved the paper. This work was funded by a research contract of the corresponding author with the Regional Health Agency (ARS) of Mayotte. We would also like to thank Julien Balicchi of the ARS of Mayotte for his contributions to data collection.

## Declarations

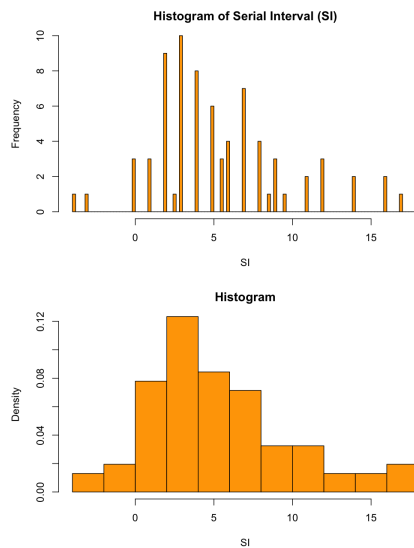
The authors declare that they have no competing interests. The co-authors have seen and agree with the content of the manuscript. We certify that the submitted work is original and is not under consideration for publication elsewhere.

## 6 Appendix: Additional tables and figures

We perform some goodness-of-fit test that determines whether or not sample data have skewness and kurtosis that matches a normal distribution.



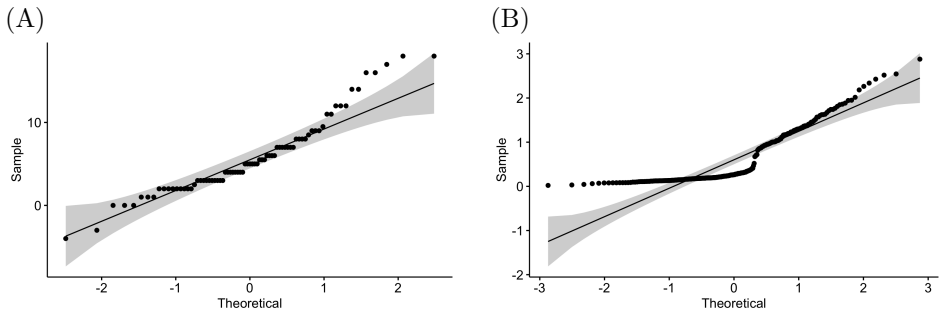
**Fig. 16** Daily reported cases from 13 March 2020 to 11 January 2022



**Fig. 17** Serial interval data distribution from 77 infector-infectee transmission pairs [37]

**Table 19** Goodness-of-fit test (p-value) of the Shapiro Wilk, Jarque-Bera, Kolmogorov-Smirnov and Anderson-Darling Normality Tests

Dataset	Shapiro Wilk	Jarque-Bera	Kolmogorov-Smirnov	Anderson-Darling
SI data	0.0005387	0.006	$< 2.2e^{-16}$	$9.297e^{-5}$
N-acetyltransferase data	$< 2.2e^{-16}$	$2.776e^{-14}$	$< 2.776e^{-16}$	$< 2.776e^{-16}$



**Fig. 18** QQ-Plot of the SI data set distribution (A) and the distribution of the N-acetyltransferase activity data set (B).

**Table 20** Skewness and Kurtosis measures

Dataset	Skewness	Kurtosis
SI dataset	0.8401043	3.57987
N-acetyltransferase data	1.197951	3.612564

## References

- [1] Samoradnitsky, G.: *Stable non-Gaussian Random Processes: Stochastic Models with Infinite Variance*. Chapman & Hall (2017)
- [2] Nolan, J.P.: *Modeling with stable distributions. Univariate Stable Distributions: Models for Heavy Tailed Data*, 25–52 (2020)
- [3] DuMouchel, W.H.: *Stable distributions in statistical inference: 2. information from stably distributed samples*. *Journal of the American Statistical Association* **70**(350), 386–393 (1975)
- [4] Zolotarev, V.M.: *One dimensional stable distributions*. American Mathematical Society (1986)
- [5] Nolan, J.P.: *Parameterizations and modes of stable distributions*. *Statistics and Probability Letters* **38**(2), 178–195 (1998)
- [6] Kanter, M.: *Stable densities under change of scale and total variation inequalities*. *The Annals of Probability* **3**(4), 697–707 (1975)
- [7] Chambers, J.M., Mallows, C.L., Stuck, B.W.: *A method for simulating stable random variables*. *Journal of the American Statistical Association* **71**(354), 340–344 (1976)
- [8] Wuertz, D., Maechler, M., Maechler, M.M.: *Package ‘stabledist’*. URL <https://cran.r-project.org/web/packages/stabledist/stabledist.pdf> (2016)

- [9] Teimouri, M., Mohammadpour, A., Nadarajah, S., Teimouri, M.M., Matrix, S., fBasics, F., RUnit, R.: Package ‘alphastable’ (2022)
- [10] Robinson, G.: Package ‘finstable’. URL <https://cran.r-project.org/web/packages/FMStable/FMStable.pdf> (2022)
- [11] Kuruoglu, E.E.: Density parameter estimation of skewed  $\alpha$ -stable distributions. *IEEE Transactions on signal processing* **49**(10), 2192–2201 (2001)
- [12] McCulloch, J.H.: Simple consistent estimators of stable distribution parameters. *Communications in statistics-simulation and computation* **15**(4), 1109–1136 (1986)
- [13] Castillo-Barnes, D., Martínez-Murcia, F.J., Ramírez, J., Górriz, J., Salas-Gonzalez, D.: Expectation–maximization algorithm for finite mixture of  $\alpha$ -stable distributions. *Neurocomputing* **413**, 210–216 (2020)
- [14] Teimouri, M.: Finite mixture of skewed sub-gaussian stable distributions. *arXiv preprint arXiv:2205.14067* (2022)
- [15] Sheather, S.J., Jones, M.C.: A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society: Series B (Methodological)* **53**(3), 683–690 (1991)
- [16] Slaoui, Y.: Bandwidth selection for recursive kernel density estimators defined by stochastic approximation method. *Journal of Probability and Statistics* **2014** (2014)
- [17] Slaoui, Y.: The stochastic approximation method for the estimation of a distribution function. *Mathematical Methods of Statistics* **23**(4), 306–325 (2014)
- [18] Salas-Gonzalez, D., Kuruoglu, E.E., Ruiz, D.P.: Finite mixture of  $\alpha$ -stable distributions. *Digital Signal Processing* **19**(2), 250–264 (2009)
- [19] Bechtel, Y.C., Bonaiti-Pellie, C., Poisson, N., Magnette, J., Bechtel, P.R.: A population and family study n-acetyltransferase using caffeine urinary metabolites. *Clinical Pharmacology & Therapeutics* **54**(2), 134–141 (1993)
- [20] Paczek, K., Jelito, D., Pitera, M., Wyłomańska, A.: Estimation of stability index for symmetric *alpha*-stable distribution using quantile conditional variance ratios. *arXiv preprint arXiv:2212.13502* (2022)
- [21] del Val, J.R., Simmross-Wattenberg, F., López, C.A., Rudis, B., Swihart, B., Rcpp, I., Rcpp, L., GSL, R.S.G., Swihart, M.B.: Package ‘libstable4u’.



- Communications in Statistics-Simulation and Computation **15**(4), 1109–1136 (2023)
- [22] Fama, E.F., Roll, R.: Parameter estimates for symmetric stable distributions. *Journal of the American statistical association* **66**(334), 331–338 (1971)
- [23] Kogon, S.M., Williams, D.B.: Characteristic function based estimation of stable distribution parameters. *A practical guide to heavy tails: statistical techniques and applications*, 311–338 (1998)
- [24] Thakur, G., Saini, J.K.: Comparative study of iterative methods for solving non-linear equations. *journal of university of shanghai for science and technology*. *Journal of University of Shanghai for Science and Technology* (2021)
- [25] Teimouri, M., Rezakhah, S., Mohammadpour, A.: Em algorithm for symmetric stable mixture model. *Communications in Statistics-Simulation and Computation* **47**(2), 582–604 (2018)
- [26] Teimouri, M.: Statistical inference for stable distribution using em algorithm. *arXiv preprint arXiv:1811.04565* (2018)
- [27] Manou-Abi, S.M., Slaoui, Y., Balicchi, J.: Estimation of some epidemiological parameters with the covid-19 data of mayotte. *Frontiers in Applied Mathematics and Statistics* **8**, 870080 (2022)
- [28] Sinaga, K.P., Yang, M.-S.: Unsupervised k-means clustering algorithm. *IEEE access* **8**, 80716–80727 (2020)
- [29] Buckle, D.: Bayesian inference for stable distributions. *Journal of the American Statistical Association* **90**(430), 605–613 (1995)
- [30] Lombardi, M.J.: Bayesian inference for  $\alpha$ -stable distributions: A random walk mcmc approach. *Computational Statistics & Data Analysis* **51**(5), 2688–2700 (2007)
- [31] Richardson, S., Green, P.J.: On bayesian analysis of mixtures with an unknown number of components (with discussion). *Journal of the Royal Statistical Society Series B: Statistical Methodology* **59**(4), 731–792 (1997)
- [32] McLachlan, G.J., Lee, S.X., Rathnayake, S.I.: Finite mixture models. *Annual review of statistics and its application* **6**, 355–378 (2019)
- [33] Nurminen, M., Mutanen, P.: Exact bayesian analysis of two proportions. *Scandinavian Journal of Statistics*, 67–77 (1987)

- [34] Matsui, M.: Fisher information matrix of general stable distributions close to the normal distribution. arXiv preprint math/0502559 (2005)
- [35] Nagaev, A., Shkol'nik, S.: Some properties of symmetric stable distributions close to the normal distribution. *Theory of Probability & Its Applications* **33**(1), 139–144 (1989)
- [36] Svensson, Å.: A note on generation times in epidemic models. *Mathematical biosciences* **208**(1), 300–311 (2007)
- [37] He, X., Lau, E.H., Wu, P., Deng, X., Wang, J., Hao, X., Lau, Y.C., Wong, J.Y., Guan, Y., Tan, X., *et al.*: Temporal dynamics in viral shedding and transmissibility of covid-19. *Nature medicine* **26**(5), 672–675 (2020)
- [38] Bechtel, Y.C., Bonaiti-Pellie, C., Poisson, N., Magnette, J., Bechtel, P.R.: A population and family study n-acetyltransferase using caffeine urinary metabolites. *Clinical Pharmacology & Therapeutics* **54**(2), 134–141 (1993)
- [39] Beaulieu, M.-C., Dufour, J.-M., Khalaf, L.: Exact confidence sets and goodness-of-fit methods for stable distributions. *Journal of Econometrics* **181**(1), 3–14 (2014)
- [40] Burnecki, K., Wylomanska, A., Chechkin, A.: Discriminating between light-and heavy-tailed distributions with limit theorem. *PLoS One* **10**(12), 0145604 (2015)
- [41] Kratz, M., Resnick, S.I.: The qq-estimator and heavy tails. *Stochastic Models* **12**(4), 699–724 (1996)
- [42] Pitera, M., Chechkin, A., Wylomańska, A.: Goodness-of-fit test for  $\alpha$ -stable distribution based on the quantile conditional variance statistics. *Statistical Methods & Applications* **31**(2), 387–424 (2022)
- [43] Diekmann, O., Heesterbeek, J.A.P., Metz, J.A.: On the definition and the computation of the basic reproduction ratio  $R_0$  in models for infectious diseases in heterogeneous populations. *Journal of mathematical biology* **28**, 365–382 (1990)
- [44] Wallinga, J., Lipsitch, M.: How generation intervals shape the relationship between growth rates and reproductive numbers. *Proceedings of the Royal Society B: Biological Sciences* **274**(1609), 599–604 (2007)
- [45] Obadia, T., Haneef, R., Boëlle, P.-Y.: The r0 package: a toolbox to estimate reproduction numbers for epidemic outbreaks. *BMC medical informatics and decision making* **12**(1), 1–9 (2012)

- [46] Boelle, P.-Y., Obadia, T., Obadia, M.T.: Package ‘r0’ (2015)
- [47] Demongeot, J., Oshinubi, K., Rachdi, M., Seligmann, H., Thuderoz, F., Waku, J.: Estimation of daily reproduction numbers during the covid-19 outbreak. *Computation* **9**(10), 109 (2021)
- [48] Waku, J., Oshinubi, K., Demongeot, J.: Maximal reproduction number estimation and identification of transmission rate from the first inflection point of new infectious cases waves: Covid-19 outbreak example. *Mathematics and Computers in Simulation* **198**, 47–64 (2022)
- [49] Geweke, J.: Bayesian inference in econometric models using monte carlo integration. *Econometrica: Journal of the Econometric Society*, 1317–1339 (1989)
- [50] Salas-Gonzalez, D., Kuruoglu, E.E., Ruiz, D.P.: Modelling with mixture of symmetric stable distributions using gibbs sampling. *Signal processing* **90**(3), 774–783 (2010)
- [51] Nagatsuka, H., Kamakura, T., Balakrishnan, N.: A consistent method of estimation for the three-parameter weibull distribution. *Computational Statistics & Data Analysis* **58**, 210–226 (2013)