Journal of Statistical Computation and Simulation
Vol. 76, No. 11, November 2006, 941–956

Taylor & Francis
Taylor & Francis Group

M-estimators with asymmetric influence functions: the $\mathcal{G}_A^0$ distribution case

HÉCTOR ALLENDE†, ALEJANDRO C. FRERY*‡, JORGE GALBIATI§ and LUIS PIZARRO†¶

†Departamento de Informática, Universidad Técnica Federico Santa María, Chile
‡Instituto de Computação, Universidade Federal de Alagoas, Campus A.C. Simões, BR 104 Norte km 14, Bloco 12, Tabuleiro dos Martins, 57072-970 Maceió, AL, Brazil
§Instituto de Estadística, Pontificia Universidad Católica de Valparaíso, Chile
¶Universidad Diego Portales, Facultad de Ingeniería, Escuela de Ingeniería Informática, Chile

(Revised 30 September 2003; in final form 13 September 2005)

Many applications consecrate the use of asymmetric distributions, and practical situations often require robust parametric inference. This paper presents the derivation of M-estimators with asymmetric influence functions, motivated by the $\mathcal{G}_A^0$ distribution. This law, regarded as the universal model for speckled imagery, can be highly skewed and maximum likelihood estimation can be severely hampered by small percentages of outliers. These outliers appear mainly because the hypothesis of independence and equal distribution of observations are seldom satisfied in practice; for instance, in the process of filtering, some pixels within a window frequently come from regions with different underlying distributions. Traditional robust estimation methods, on the basis of symmetric robustifying functions, assume that the distribution is symmetric, but when the data distribution is asymmetric, these methods yield biased estimators. Empirical influence functions for maximum likelihood estimators are computed, and based on this information we propose the asymmetric M-estimator (AM-estimator), an M-estimator with asymmetric redescending functions. The performance of AM estimators is assessed, and it is shown that they either compete with or outperform both maximum likelihood and Huber-type M-estimators.

Keywords: Robustness; Multiplicative Model; Speckle Noise

1. Introduction

The precise knowledge of the statistical properties data plays a central role in image processing and understanding. In remote sensing, for instance, these properties can be used to discriminate types of land use and to develop specialised filters for speckle noise reduction, among other applications. Statistical image filtering, segmentation and classification are procedures that heavily rely on dependable inference procedures [1].

*Corresponding author. Email: acfrery@pesquisador.cnpq.br

Journal of Statistical Computation and Simulation
ISSN 0094-9655 print/ISSN 1563-5163 online
http://www.tandf.co.uk/journals
DOI: 10.1080/10629360600569154

Statistical modelling is of particular relevance when dealing with speckled data. Speckle noise appears when coherent illumination is employed, as is the case of sonar, laser, ultrasound and synthetic aperture radar (SAR) imagery. This kind of degradation severely impairs human and machine ability to discriminate targets, and it is known to defy the classical assumptions of additivity and Gaussian distribution.

There are many statistical models for speckled data [2]. Among them, the multiplicative model is based on the assumption that the observed random field Z is the result of the product of two independent and unobserved random fields: X and Y . The random field X models the terrain backscatter, and thus depends only on the type of area each pixel belongs to. The random field Y describes the speckle noise, taking into account that L (ideally) independent images are averaged in order to reduce the noise.

There are various ways of modelling the random fields X, whereas the physics of speckle noise allows the assumption of a $\Gamma^{1/2}$ law for Y. The universal model [3, 4] proposes the $\Gamma^{-1/2}$ distribution to describe the amplitude backscatter X, yielding the $\mathcal{G}_A^0$ distribution for the return. One of the advantages of the $\mathcal{G}_A^0$ distribution over the classical K_A distribution is that it models very well extremely heterogeneous areas like cities, as well as moderately heterogeneous areas like forests and homogeneous areas like crops. This law has also been used to describe different types of tissue in B-scan ultrasound imagery [5].

The $\mathcal{G}_A^0$ distribution is characterised by as many parameters as the K_A distribution: the number of looks (L), the scale parameter (γ) and the roughness parameter (α). This distribution has the same nice interpretational properties that the K_A distribution has. The parameter α is of particular interest in many applications, as it is directly related to the roughness of the target, and γ is a scale parameter related to the relative power between reflected and incident signals.

This work discusses the problem of estimating the parameters of the $\mathcal{G}_A^0$ for the single-look case. This is the noisiest case and, therefore, images with $L = 1$ are the ones that depend more on reliable inference procedures for, for instance, filtering [6, 7] and classification [8, 9].

Robustness is a desirable property for estimators, as it allows their use even in situations where the quality of the input data is unreliable. A situation where this occurs is when ground control points appear in the SAR image, which are essential for data calibration. These points produce a higher return than the rest of the image, and for this reason they are called 'corner reflectors'. If data from corner reflectors are included in an analysis with non-robust estimation procedures, the results may be completely unreliable, as they behave as outliers in the sample. Another typical situation arises when applying filters [10]; data are collected in a window, and there is no way to assure that they form a perfect uncontaminated sample.

When the distribution is symmetric, the problems caused by outliers can be reduced using 'classical' robust estimation techniques [11], which tend to ignore or put less weight on influential observations on both sides of the mean.

In many applications, one often finds data distributions with asymmetric heavy tails [12], as the $\mathcal{G}_A^0$ distribution. Dealing with such data is essentially difficult because samples from the tail of the distribution will have a strong influence on parameter estimation, and downweighting them introduces bias [13].

Section 2 of this paper reviews the fundamental properties of the model considered here, with results regarding the use of Stylised Empirical Influence Functions for estimators assessment, and presents the main estimation techniques available for the $\mathcal{G}_A^0$ distribution. Section 3 introduces the robust estimators, classical M-estimators and the novel AM-estimators. Section 4 shows a comparative simulation study of robust estimators. In section 5, concluding remarks and future extensions are presented.

2. The model

The general (multilook) form of the density which characterises the $\mathcal{G}_A^0\ (\alpha, \gamma, L)$ distribution is given by

$$f(x) = \frac{2L^L\Gamma(L-\alpha)}{\gamma^\alpha\Gamma(L)\Gamma(-\alpha)}\frac{x^{2L-1}}{(\gamma+Lx^2)^{L-\alpha}}, \quad x > 0 \tag{1}$$

where $\alpha < 0$ is referred to as the roughness parameter, $\gamma > 0$ is the scale parameter and $L \geq 1$ is the number of looks. The number of looks is controlled in the early generation steps of the image, and is known beforehand or it is estimated using extended homogeneous targets; this parameter remains constant over all the images.

This law was originally devised to describe extremely heterogeneous clutter [3], and lately proposed and assessed as a universal model for speckled imagery [4]. Improved estimation using bootstrap for the parameters α and γ of this distribution is presented by Cribari-Neto *et al.* [14], whereas the robustness for the $L = 1$ case is studied by Bustos *et al.* [15] using Huber-type M-estimators.

As commented by Bustos *et al.* [15], the single-look case is of particular interest, and it will be considered here, as it describes the noisiest images. The distribution of interest is, then, characterised by the density

$$f(x; (\alpha, \gamma)) = -\frac{2\alpha}{\gamma^\alpha}\frac{x}{(\gamma+x^2)^{1-\alpha}} = \frac{2\alpha x}{\gamma(1+x^2/\gamma)^{1-\alpha}}, \quad x > 0 \tag{2}$$

with $-\alpha, \gamma > 0$. This distribution will be denoted $\mathcal{G}_A^0(\alpha, \gamma)$. Its cumulative distribution function is given by

$$F(x; (\alpha, \gamma)) = 1 - (1 + x^2/\gamma)^\alpha, \tag{3}$$

and its inverse, useful for the generation of random deviates and the computation of quantiles, is given by

$$F^{-1}(t) = \sqrt{\gamma((1-t)^{1/\alpha} - 1)}. \tag{4}$$

Several parameter estimation techniques are available to estimate $\theta = (\alpha, \gamma)$, being the most remarkable ones those based on sample moments and maximum likelihood. The kth order moment of the $\mathcal{G}_A^0(\alpha, \gamma)$ distribution is given by

$$E(X^k) = \begin{cases} \gamma^{k/2}\dfrac{k\Gamma(k/2)\Gamma(-\alpha-k/2)}{2\Gamma(-\alpha)} & \text{if } -\alpha > k/2 \\ \infty & \text{else.} \end{cases} \tag{5}$$

Denoting the jth sample moment $m_j = N^{-1}\sum_{i=1}^N x_i^j$ and using equation (5), it is possible to compute the MO-estimators $\hat\theta_{\mathrm{MO}}$ by means of the half and first-order moments:

$$\begin{cases} m_{1/2} = \hat\gamma_{\mathrm{MO}}^{1/4}\dfrac{\Gamma(-\hat\alpha_{\mathrm{MO}}-1/4)\Gamma(1/4)}{4\Gamma(-\hat\alpha_{\mathrm{MO}})} \\ m_1 = \hat\gamma_{\mathrm{MO}}^{1/2}\dfrac{\sqrt{\pi}\Gamma(-\hat\alpha_{\mathrm{MO}}-1/2)}{2\Gamma(-\hat\alpha_{\mathrm{MO}})}, \end{cases} \tag{6}$$

assuming that $-\alpha > 1/2$.

The maximum likelihood estimator, based on $x_1, \dots, x_N$, is defined as a value $\hat{\theta}_{\mathrm{ML}}$ which minimises $-\sum_{i=1}^{N} \ln f(x_i;\theta)$. Equating to zero the derivates of this function, we obtain

$$\sum_{i=1}^{N} s(x_i;\theta) = 0, \tag{7}$$

where

$$s(x;\theta) = (s_1(x;\theta), s_2(x;\theta))^{\mathrm{T}} = \frac{\partial}{\partial\theta} \ln f_\theta(x) = \left(\frac{\partial}{\partial\theta_1} \ln f_\theta(x), \frac{\partial}{\partial\theta_2} \ln f_\theta(x)\right)^{\mathrm{T}}$$

denotes the vector of likelihood scores. In our case

$$\begin{cases} s_1(x;\theta) = \dfrac{1}{\alpha} + \ln\left(1 + \dfrac{x^2}{\gamma}\right) \\ s_2(x;\theta) = \dfrac{-\alpha}{\gamma} - \dfrac{1-\alpha}{\gamma - x^2} \end{cases} \tag{8}$$

From here, we derive the estimator $\hat{\theta}_{\mathrm{ML}} = (\hat{\alpha}_{\mathrm{ML}}, \hat{\gamma}_{\mathrm{ML}})$ as:

$$\begin{aligned} \hat{\alpha}_{\mathrm{ML}} &= -\left(\frac{1}{N}\sum_{i=1}^{N} \ln\left(1 + \frac{x_i^2}{\hat{\gamma}_{\mathrm{ML}}}\right)\right) \\ \hat{\gamma}_{\mathrm{ML}} &= \left[\left(1 + \frac{1}{N}\sum_{i=1}^{N} \ln\left(1 + \frac{x_i^2}{\hat{\gamma}_{\mathrm{ML}}}\right)\right) \frac{1}{N}\sum_{i=1}^{N} \left(\hat{\gamma}_{\mathrm{ML}} + x_i^2\right)^{-1}\right]^{-1}. \end{aligned} \tag{9}$$

Maximum likelihood estimation for the K_A distribution was considered in the work by Joughin *et al.* [16].

3. Robust estimators

Robust estimation has become more prevalent in remote sensing with the emergence of a new generation of sensors. While the new sensor technology provides higher spectral and spatial resolution, enabling a greater number of spectrally separable classes to be identified, labelled samples for designing the classifier remains difficult and expensive.

Outliers are not uncommon in the practice of image analysis, where scenes usually contain pixels of unknown origin. The statistical distribution of these pixels may be significantly different from the training classes and can constitute statistical outliers. Unfortunately, these outlying pixels are usually scattered throughout the scene and are small in number and, therefore, identifying these pixels could be a tedious task. A common approach to eliminate the effect of those pixels is to use robust techniques [1, 8].

3.1 *Traditional M-estimators*

Let us first consider the problem of estimating from a finite sample the unconditional mean of a heavy-tailed distribution. The empirical average may be a poor estimator here because a few points will be sampled from the tails and may have very different values, thereby introducing a great deal of variability in the empirical average. In the case of a symmetric distribution, we can downweight or ignore the effect of these outliers in order to reduce variability.

One of the best known classes of robust estimators are M-estimators, a class of bounded-influence estimators, which are a generalisation of the ML-estimators. Consider a set of data $x_1, \ldots, x_N$, independent samples from the same symmetric distribution (symmetric i.i.d) with parameter θ. The M-estimate $\hat{\theta}_{\mathrm{M}}$ is defined as the minimum of a global energy function

$$\hat{\theta}_{\mathrm{M}} = \arg\min_{\theta} E(\theta), \tag{10}$$

where the energy function $E(\theta)$ is defined in terms of a loss function ρ as

$$E(\theta) = \sum_{i=1}^{N} \rho(x_i; \theta). \tag{11}$$

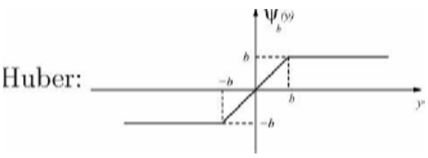
Equivalently, one solves the estimation equation

$$\sum_{i=1}^{N} \psi(x_i; \theta) = 0, \tag{12}$$

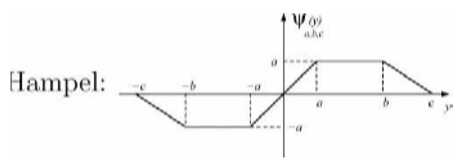
where $\psi(x_i; \theta) = \partial\rho(x_i; \theta)/\partial\theta$.

Typically, the function ρ is chosen symmetric; particular cases are $\rho(y) = y^2/2$, yielding the least square (LS) estimator, and $\rho(y) = |y|$ yielding the median estimator. Equation (12) is a generalisation of equation (7), as ML-estimators are obtained considering $\rho(x; \theta) = -\ln f_\theta(x)$ and $\psi(x; \theta) = s(x; \theta)$.

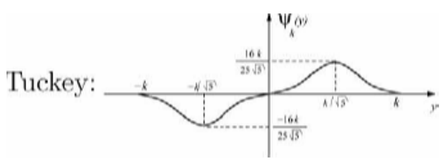
The robustifying functions ψ are a composition of score functions and bounded symmetric functions, usually defined by parts, for instance:



$$\psi_b(y) =\in \{b, \max\{y, -b\}\} \tag{13}$$



$$\psi_{a,b,c}(y) = \begin{cases} y & 0 \le |y| \le a \\ a \cdot \mathrm{sign}(y) & a \le |y| \le b \\ a\dfrac{c - |y|}{c - b}\mathrm{sign}(y) & b \le |y| \le c \\ 0 & c \le |y| \end{cases} \tag{14}$$



$$\psi_k(y) = \begin{cases} y\left(1 - \left(\dfrac{y}{k}\right)^2\right)^2 & |y| \le k \\ 0 & |y| > k \end{cases} \tag{15}$$

The tuning parameters a, b, c, k are obtained requiring that the asymptotic relative efficiency of the M-estimator, with respect to the ML-estimator in the model without outliers, ranges from, for instance, 90–95%.

In order to obtain unbiased and optimal estimators, we redefine the M-estimator $\hat{\theta}_{\mathrm{M}}$ as a solution of the equation

$$\sum_{i=1}^{N} \psi[s(x_i;\theta) - c] = 0 \tag{16}$$

where the Fisher consistency is accomplished by means of the c function, which is defined implicitly as

$$\int_{-\infty}^{\infty} \psi[s(x_i;\theta) - c]\, \mathrm{d}F_\theta(x) = 0 \tag{17}$$

Many theoretical results concerning the asymptotic and the robustness properties of M-estimators are available in the literature [11, 17].

In the following, we will see the need to use asymmetric robustifying functions to deal with speckled data.

3.2 *Asymmetric M-estimators*

A qualitative way to describe the robustness of the estimators is by the empirical influence function (EIF) [11]. The EIF shows what happens with the estimator T_N when an observation x ranges over the support of the distribution. It is defined as $\mathrm{EIF}(x) = T_N(x_1, x_2, \ldots, x_{N-1}, x)$. In order to make the value of EIF(x) independent of the particular sample, we will use the stylised empirical influence function (SEIF) proposed by Andrews *et al.* [18], which consists of using the ith quantile of the underlying distribution

$$x_i = F^{-1}\left(\frac{i - 1/3}{N + 1/3}\right). \tag{18}$$

When we work with symmetric distributions, their SEIF are symmetric too; because of this reason, typically symmetric robustifying functions are selected [19]. In our case, the $\mathcal{G}_A^0$ is a non-symmetric distribution, so is the SEIF of its ML-estimator, as shown in figure 1.

In figure 1, we note that the loss of robustness depends on the size of the outlier x and on the type of area (true α). In homogeneous areas ($\alpha = -10$, lower right), it tends to be critical for large values of the outlier, whereas in extremely heterogeneous areas ($\alpha = -1$, upper left), this occurs for small ones. Regarding heterogeneous areas (upper right and lower left), the loss of robustness oscillates between both extremes. Besides that, and independently of the roughness, if the sample size decreases, the asymmetry of the SEIF becomes more pronounced.

To overcome this issue, we propose the use of a family of non-symmetric redescending robustifying functions Ψ_{r_1,r_2}, where $0 < r_1, r_2 < \infty$ are the tuning parameters: all the asymmetric functions that are zero outside the interval $[r_1, r_2]$ and that satisfy the general conditions presented by Hampel *et al.* [11, p. 126].

Redescending asymmetric piecewise linear functions belong to this family for either $r_1 \le r_2$ or $r_1 \ge r_2$. These functions are given by

$$\psi_{r_1,r_2}(y) = \begin{cases} -y - r_1 & -r_1 \le y < -r_1/2 \\ y & -r_1/2 \le y < r_2/2 \\ -y + r_2 & r_2/2 \le y \le r_2 \\ 0 & \text{else,} \end{cases} \tag{19}$$

see figure 2.

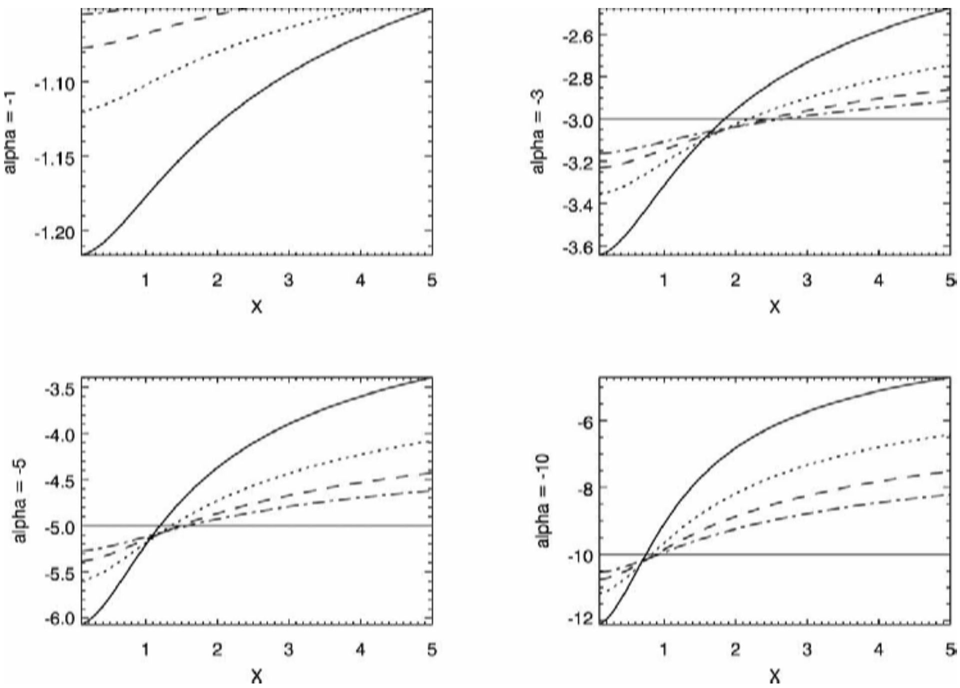


Figure 1. Stylised empirical influence functions for the ML-estimator under the $\mathcal{G}_A^0(\alpha, 1)$ distribution with $-\alpha \in \{1, 3, 5, 10\}$ for $N = 9$ (solid line), $N = 25$ (dashed line), $N = 49$ (long-dashed line) and $N = 81$ (dashed long-dashed line).

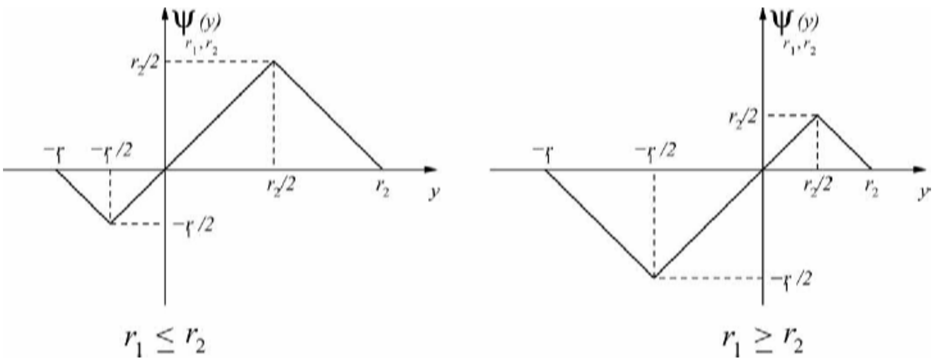


Figure 2. Asymmetric robustifying redescending piecewise linear function.

Then, an AM-estimator $\hat{\theta}_{\mathrm{AM}}$ is defined as in equation (16), but with an asymmetric function of type $\psi_{r_1,r_2} \in \Psi_{r_1,r_2}$. The choice of r_1 and r_2, as well as c, will be discussed in the next section.

4. Simulation study

In order to assess the behaviour of estimators robustified with Ψ_{r_1,r_2} functions (AM-estimators), we consider a model without contamination and several models with outliers. The pattern of contamination is defined as a sequence of i.i.d. random variables $x_1, \ldots, x_N$, with common distribution function

$$F(x; (\alpha, \gamma); \epsilon; z_v) = (1 - \epsilon)F(x; (\alpha, \gamma)) + \varepsilon\delta_{z_v}(x) \tag{20}$$

where $\delta_{z_v}(x) = 1_{[z_v;+\infty)}(x)$ with z_v a very large value when compared with most of the sample data: a factor of the sample mean $z_v = vE[X]$, and $\epsilon \in [0, 1]$ the probability that an observation is an outlier. Therefore, in a sample of N data, we will have on average $(1 - \epsilon)N$ data with distribution $\mathcal{G}_A^0$ and ϵN outliers. The samples were generated using relation (4), choosing the scale parameter γ in terms of α, so that $E[X] = 1$. Using equation (6) we have

$$\gamma = \gamma_\alpha = \frac{4}{\pi}\left[\frac{\Gamma(-\alpha)}{\Gamma(-\alpha - 1/2)}\right].$$

In this work, we compare the AM-estimators of the distribution $\mathcal{G}_A^0$, using the robustifying function described in equation (19), with respect to the ML and M-estimators, the latter based on the Huber function, equation (13), presented by Bustos *et al.* [15].

For simplicity, the tuning parameters r_1 and r_2 are made to depend on one another and are related with the tuning parameter b of the Huber function in the following ways: case (i) $r_1 = 2b$ and $r_2 = \eta r_1$; case (ii) $r_2 = 2b$ and $r_1 = \eta r_2$, varying $b = \{1/2, 1, 2, 3, 4\}$ and $\eta = \{1/2, 1, 3/2, 2\}$, where η is used to control the amplitude of the asymmetry of the robustifying function. The estimators were implemented as suggested by Marazzi and Ruffieux [20].

The estimators were compared by means of a Monte Carlo experience. The mean, mean square error and absolute relative bias were estimated using $R = 1000$ replications and observing, respectively, $\hat{E}[\hat{\theta}] = R^{-1}\sum_{i=1}^{R}\hat{\theta}_i$, $\widehat{\mathrm{MSE}} = \hat{E}[\hat{\theta} - \theta]^2 = \hat{V}(\hat{\theta}) + (\hat{E}[\hat{\theta}] - \theta)^2$, and $\hat{B}[\hat{\theta}] = \theta^{-1}|\hat{E}[\hat{\theta}] - \theta|$, where θ is the true value of the parameter and $\hat{\theta}$ is the estimator.

Figures 3 and 4 show a graphical comparison of estimators for $\alpha = \{-1, -10\}$, changing both the probability $\epsilon = \{0.01, 0.05, 0.10\}$ and the magnitude $v = \{5, 10, 20, 40\}$ of the contamination.

The first fact that stands out in the results is that AM-estimators of type (ii), namely those with $r_2 = 2b$ and $r_1 = \eta r_2$, exhibit similar behaviour in all the cases, indicating that for ψ_{r_1,r_2} in $\mathbb{R}^-$, it is possible to work with a single parametrisation, say $\eta = 1$, for all α. This happens because of the aforementioned behaviour of the SEIF of the ML-estimator: its negative bias is small and the loss of robustness tends to be critical for large values of the outliers (figure 1). Therefore, it is sufficient to study the behaviour of AM-estimators of type (i), *i.e.* with $r_1 = 2b$ and $r_2 = \eta r_1$.

It is also clear that as data lose homogeneity, the Huber M-estimator gains precision, that is to say, in homogeneous areas it is less precise than in heterogeneous areas, becoming less precise as the proportion and magnitude of the contamination increase, because of the constant weight that the function assigns to extreme values. In contrast, AM-estimators achieve good precision independently of the homogeneity, especially as the proportion and magnitude of the contamination increase because of the decreasing weight that the function ψ_{r_1,r_2} assigns to extreme values.

As for the asymmetry of the function ψ_{r_1,r_2}, we observe that it yields better results for homogeneous areas when $\eta = \{1/2, 1\}$, whereas for extremely heterogeneous areas, this

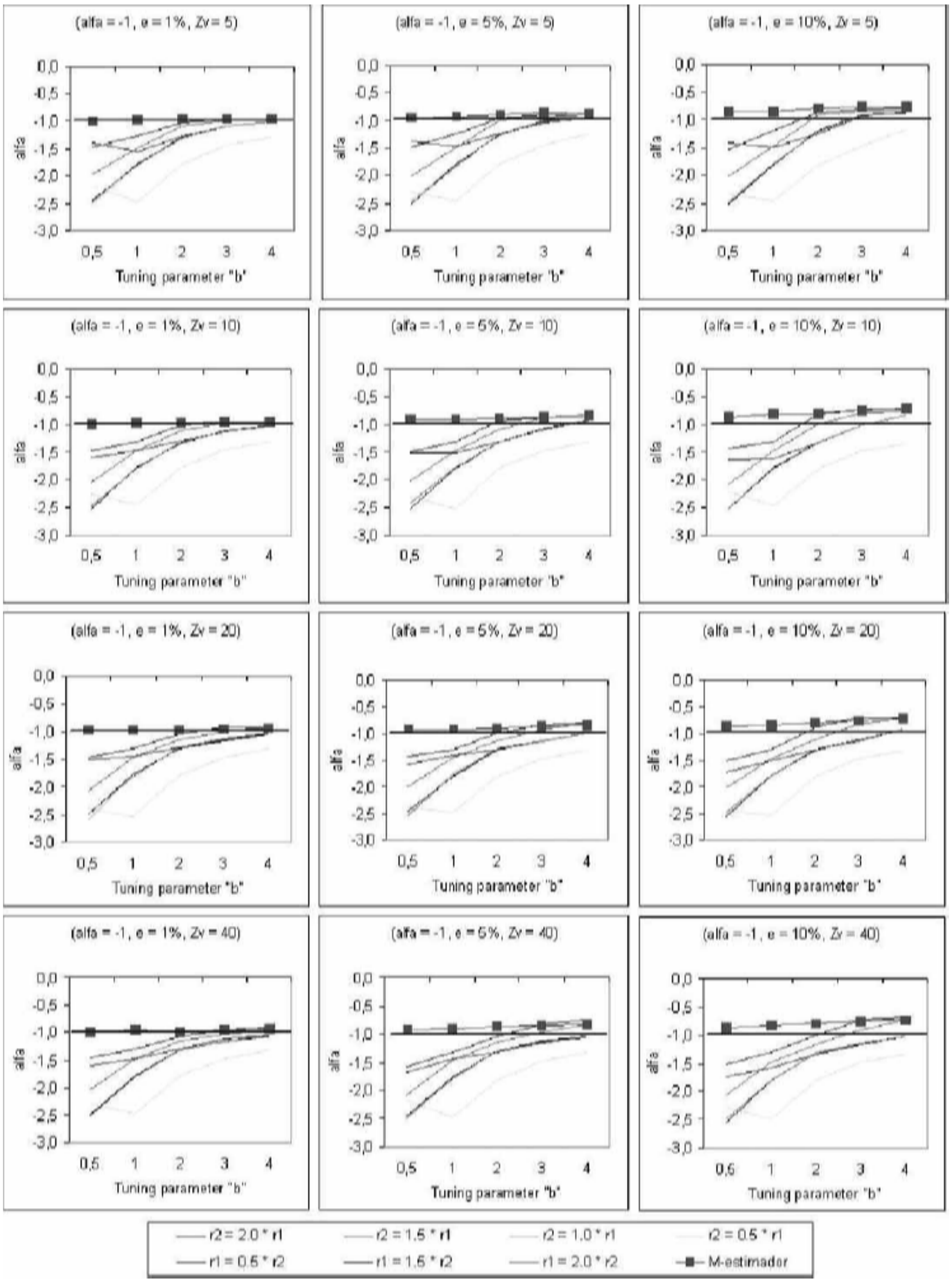


Figure 3. AM-estimators versus M-estimator for $\alpha = -1$.

occurs when $\eta = \{3/2, 2\}$, and in heterogeneous areas all the parametrisations present a similar behaviour. At this point, it is interesting to relate the properties of the $\mathcal{G}_A^0$ distribution with the asymmetry of the robustifying function, in order to obtain the most adequate parametrisation for each level of contamination; we could appeal to kth order moment and other robust centre and dispersion measures.

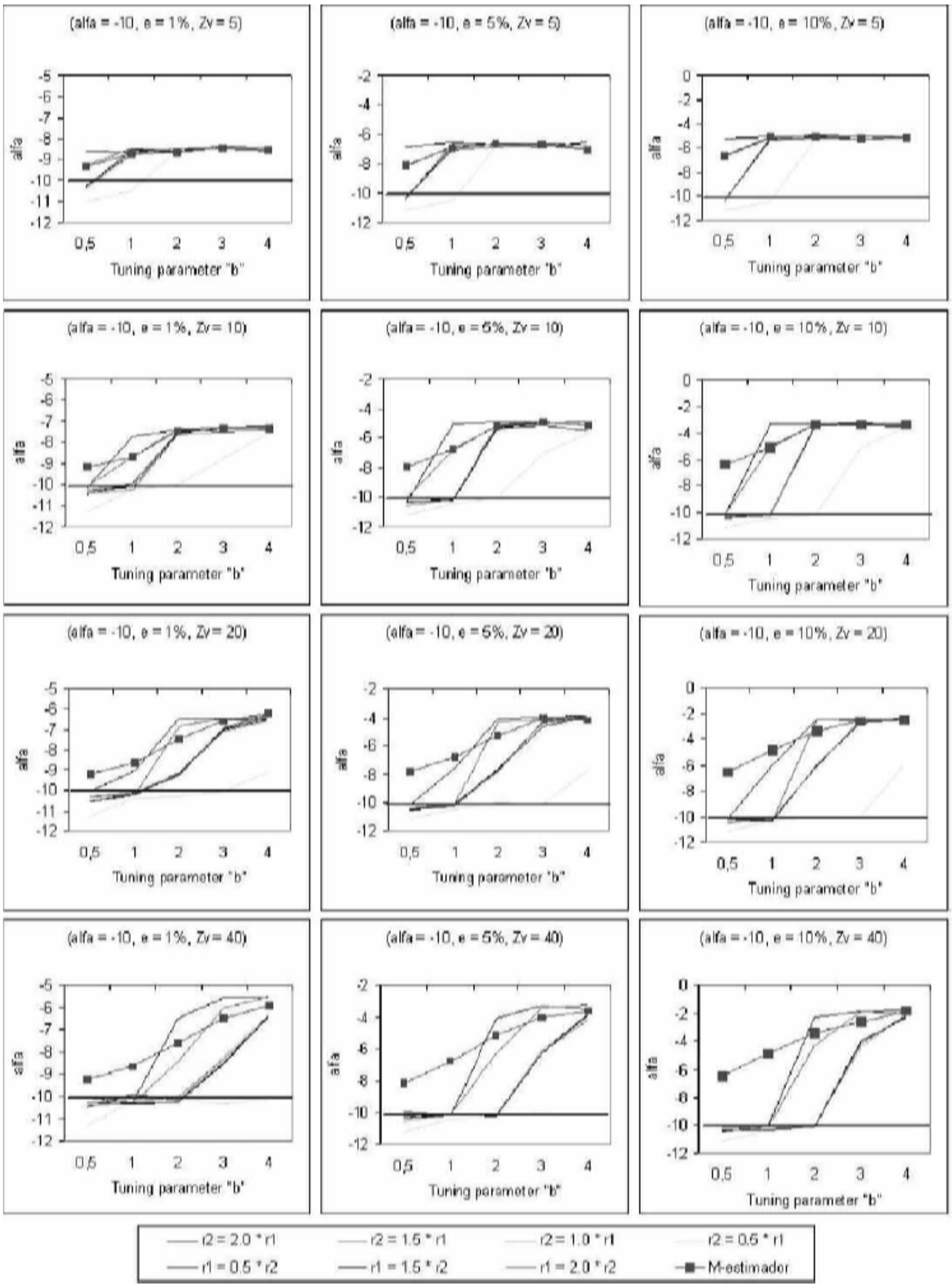


Figure 4. AM-estimators versus M-estimator for $\alpha = -10$.

The rule to choose the tuning parameter b and, consequently, r_1 and r_2, consists of requiring that the asymptotic relative efficiency (ARE) of the AM-estimator with respect to the ML-estimator, in the model without outliers, satisfies $90\% \leq \text{ARE} \leq 95\%$.

Tables 1–6 show the quantitative comparison among ML, M and AM-estimators. The tuning parameters of the M-estimator were also chosen to satisfy the efficiency criterion.

Table 1. Numerical comparison of the mean among ML, M and AM estimators, for varying α, sample size and contamination level ϵ, with $v = 5$ (winning cases in boldface).

		$\alpha=-1$			$\alpha=-6$			$\alpha=-10$		
ϵ	N	$\hat{E}[\hat{\alpha}_{ML}]$	$\hat{E}[\hat{\alpha}_{M}]$	$\hat{E}[\hat{\alpha}_{AM}]$	$\hat{E}[\hat{\alpha}_{ML}]$	$\hat{E}[\hat{\alpha}_{M}]$	$\hat{E}[\hat{\alpha}_{AM}]$	$\hat{E}[\hat{\alpha}_{ML}]$	$\hat{E}[\hat{\alpha}_{M}]$	$\hat{E}[\hat{\alpha}_{AM}]$
0.00	9	−1.162	**−1.140**	−1.154	−6.520	−6.519	**−6.517**	−9.916	−9.916	**−9.905**
	25	−1.048	**−1.041**	−1.045	−6.283	**−6.282**	**−6.282**	−10.309	−10.309	**−10.304**
	49	−1.013	−1.004	**−1.002**	**−6.117**	**−6.117**	−6.118	−10.174	−10.174	**−10.173**
	81	−1.014	**−1.012**	−1.013	**−6.065**	**−6.065**	**−6.065**	−10.114	−10.114	**−10.112**
0.01	9	−0.826	−0.920	**−0.982**	−3.143	−3.130	**−3.116**	−4.655	−4.653	**−6.173**
	25	−0.911	−0.943	**−0.958**	**−4.501**	−4.499	−4.496	−6.976	−6.976	**−8.371**
	49	−0.940	−0.957	**−0.964**	**−4.993**	−4.992	−4.991	−8.080	−8.080	**−9.065**
	81	−0.956	−0.967	**−0.971**	**−5.266**	−5.265	−5.264	−8.605	−8.605	**−9.360**
0.05	9	−0.815	−0.909	**−0.965**	**−2.956**	−2.939	−2.920	−4.385	−4.380	**−5.696**
	25	−0.858	−0.900	**−0.918**	**−3.993**	−3.987	−3.980	−6.127	−6.126	**−7.546**
	49	−0.872	−0.905	**−0.914**	**−4.299**	−4.295	−4.291	−6.682	−6.682	**−7.989**
	81	−0.875	−0.908	**−0.915**	**−4.342**	−4.338	−4.335	−6.715	−6.715	**−8.010**
0.10	9	−0.792	−0.889	**−0.950**	**−2.732**	−2.708	−2.682	−4.011	−4.003	**−5.048**
	25	−0.800	−0.848	**−0.863**	**−3.366**	−3.354	−3.341	−5.097	−5.094	**−6.366**
	49	−0.785	−0.827	**−0.832**	**−3.446**	−3.436	−3.426	−5.156	−5.155	**−6.472**
	81	−0.774	−0.815	**−0.817**	**−3.362**	−3.352	−3.342	−5.101	−5.099	**−6.446**

Table 2. Numerical comparison of the mean among ML, M and AM estimators, for varying α, sample size and contamination level ϵ, with $v = 15$ (winning cases in boldface).

		$\alpha=-1$			$\alpha=-6$			$\alpha=-10$		
ϵ	N	$\hat{E}[\hat{\alpha}_{ML}]$	$\hat{E}[\hat{\alpha}_{M}]$	$\hat{E}[\hat{\alpha}_{AM}]$	$\hat{E}[\hat{\alpha}_{ML}]$	$\hat{E}[\hat{\alpha}_{M}]$	$\hat{E}[\hat{\alpha}_{AM}]$	$\hat{E}[\hat{\alpha}_{ML}]$	$\hat{E}[\hat{\alpha}_{M}]$	$\hat{E}[\hat{\alpha}_{AM}]$
0.00	9	−1.162	**−1.140**	−1.154	−6.508	−6.507	**−6.506**	−9.997	−9.997	**−9.986**
	25	−1.048	**−1.041**	−1.045	−6.265	**−6.264**	**−6.264**	−10.295	−10.295	**−10.290**
	49	−1.013	−1.004	**−1.002**	**−6.114**	**−6.114**	**−6.114**	−10.175	−10.175	**−10.174**
	81	−1.014	**−1.012**	−1.013	**−6.060**	**−6.060**	**−6.060**	−10.123	−10.123	**−10.122**
0.01	9	−0.682	**−0.920**	−1.258	−1.818	−2.801	**−6.657**	−2.298	−3.343	**−9.839**
	25	−0.837	−0.943	**−1.044**	−3.245	−4.355	**−6.293**	−4.432	−5.957	**−10.211**
	49	−0.894	−0.957	**−1.013**	−4.042	−4.937	**−6.125**	−5.961	−7.379	**−10.242**
	81	−0.922	−0.967	**−1.006**	−4.464	−5.190	**−6.049**	−6.808	−8.036	**−10.194**
0.05	9	−0.668	**−0.909**	−1.285	−1.691	−2.592	**−6.607**	−2.130	−3.080	**−9.938**
	25	−0.767	−0.900	**−1.052**	−2.701	−3.787	**−6.371**	−3.695	−5.080	**−10.208**
	49	−0.796	−0.905	**−1.022**	−3.112	−4.146	**−6.163**	−4.286	−5.700	**−10.228**
	81	−0.802	−0.908	**−1.016**	−3.156	−4.183	**−6.110**	−4.346	−5.771	**−10.148**
0.10	9	−0.638	**−0.886**	−1.354	−1.553	−2.365	**−6.655**	−1.957	−2.798	**−9.949**
	25	−0.701	−0.861	**−1.085**	−2.147	−3.111	**−6.329**	−2.835	−3.975	**−10.070**
	49	−0.681	−0.830	**−1.034**	−2.136	−3.110	**−6.128**	−2.877	−4.066	**−10.199**
	81	−0.666	−0.814	**−1.019**	−2.068	−3.052	**−6.072**	−2.752	−3.941	**−10.149**

The results in the tables show that all ML, M and AM-estimators exhibit almost the same behaviour when the sample is free of contamination. Besides, when the sample size grows, all methods show better estimates. Nevertheless, when the percentage of outliers increases, the ML and M-estimators lose accuracy faster than the AM-estimators. Summarising, the AM-estimators show either the same or better performance than ML and M-estimators in all cases.

Table 3. Numerical comparison of the mean among ML, M and AM estimators, for varying α, sample size and contamination level ϵ, with $v = 15$ (winning cases in boldface).

		$\alpha = -1$			$\alpha = -6$			$\alpha = -10$		
ϵ	N	$\widehat{\mathrm{mse}}[\hat{\alpha}_{\mathrm{ML}}]$	$\widehat{\mathrm{mse}}[\hat{\alpha}_{\mathrm{M}}]$	$\widehat{\mathrm{mse}}[\hat{\alpha}_{\mathrm{AM}}]$	$\widehat{\mathrm{mse}}[\hat{\alpha}_{\mathrm{ML}}]$	$\widehat{\mathrm{mse}}[\hat{\alpha}_{\mathrm{M}}]$	$\widehat{\mathrm{mse}}[\hat{\alpha}_{\mathrm{AM}}]$	$\widehat{\mathrm{mse}}[\hat{\alpha}_{\mathrm{ML}}]$	$\widehat{\mathrm{mse}}[\hat{\alpha}_{\mathrm{M}}]$	$\widehat{\mathrm{mse}}[\hat{\alpha}_{\mathrm{AM}}]$
0.00	9	**0.218**	**0.218**	0.234	**5.265**	5.267	5.274	**5.824**	5.825	5.886
	25	**0.046**	0.052	0.058	**1.801**	1.802	1.804	**3.516**	**3.516**	3.529
	49	**0.021**	0.024	0.028	**0.799**	**0.799**	**0.799**	**2.207**	**2.207**	**2.207**
	81	**0.014**	0.016	0.018	0.477	**0.476**	**0.476**	**1.377**	**1.377**	**1.377**
0.01	9	**0.056**	0.078	0.129	**8.471**	8.555	8.645	29.109	29.140	**17.442**
	25	**0.033**	0.042	0.052	**2.775**	2.786	2.798	10.350	10.352	**5.106**
	49	**0.020**	0.024	0.029	**1.461**	1.465	1.469	4.934	4.934	**2.596**
	81	**0.012**	0.014	0.016	**0.915**	0.917	0.919	2.988	2.989	**1.573**
0.05	9	**0.058**	0.072	0.115	**9.673**	9.794	9.928	32.389	32.450	**22.094**
	25	**0.043**	0.045	0.053	**4.778**	4.812	4.846	16.932	16.941	**9.114**
	49	0.032	**0.031**	0.036	**3.550**	3.568	3.586	12.859	12.862	**6.129**
	81	0.026	**0.022**	0.024	**3.224**	3.237	3.251	12.259	12.262	**5.378**
0.10	9	**0.069**	0.078	0.125	**11.223**	11.403	11.609	37.020	37.130	**28.813**
	25	0.060	**0.053**	0.058	**7.754**	7.836	7.925	26.177	26.213	**17.022**
	49	0.058	**0.048**	0.051	**7.091**	7.152	7.217	24.965	24.988	**14.636**
	81	0.058	**0.045**	0.048	**7.308**	7.367	7.429	24.873	24.891	**13.960**

Table 4. Numerical comparison of the mean among ML, M and AM estimators, for varying α, sample size and contamination level ϵ, with $v = 15$ (winning cases in boldface).

		$\alpha = -1$			$\alpha = -6$			$\alpha = -10$		
ϵ	N	$\widehat{\mathrm{mse}}[\hat{\alpha}_{\mathrm{ML}}]$	$\widehat{\mathrm{mse}}[\hat{\alpha}_{\mathrm{M}}]$	$\widehat{\mathrm{mse}}[\hat{\alpha}_{\mathrm{AM}}]$	$\widehat{\mathrm{mse}}[\hat{\alpha}_{\mathrm{ML}}]$	$\widehat{\mathrm{mse}}[\hat{\alpha}_{\mathrm{M}}]$	$\widehat{\mathrm{mse}}[\hat{\alpha}_{\mathrm{AM}}]$	$\widehat{\mathrm{mse}}[\hat{\alpha}_{\mathrm{ML}}]$	$\widehat{\mathrm{mse}}[\hat{\alpha}_{\mathrm{M}}]$	$\widehat{\mathrm{mse}}[\hat{\alpha}_{\mathrm{AM}}]$
0.00	9	**0.218**	**0.218**	0.234	**5.316**	5.320	5.325	**6.036**	**6.036**	6.101
	25	**0.046**	0.052	0.058	**1.647**	**1.647**	1.648	**3.636**	**3.636**	3.650
	49	**0.021**	0.024	0.028	**0.782**	**0.782**	**0.782**	**2.189**	**2.189**	**2.189**
	81	**0.014**	0.016	0.018	**0.444**	**0.444**	**0.444**	**1.415**	**1.415**	1.417
0.01	9	0.114	**0.078**	0.310	17.546	10.494	**6.079**	59.377	44.509	**6.389**
	25	0.045	**0.042**	0.061	7.815	3.225	**1.946**	31.439	17.195	**3.458**
	49	0.026	**0.024**	0.030	4.219	1.637	**0.861**	17.194	8.014	**2.208**
	81	0.016	**0.014**	0.016	2.777	1.051	**0.488**	11.325	4.962	**1.461**
0.05	9	0.124	**0.072**	0.314	18.690	12.054	**5.538**	62.132	48.426	**5.854**
	25	0.074	**0.045**	0.061	11.444	5.793	**2.143**	40.882	26.000	**3.386**
	49	0.058	**0.031**	0.032	9.027	4.165	**0.862**	34.517	20.700	**2.298**
	81	0.050	0.022	**0.018**	8.690	3.835	**0.490**	33.603	19.612	**1.402**
0.10	9	0.148	**0.078**	0.480	19.958	13.810	**5.859**	64.993	52.696	**6.463**
	25	0.110	**0.052**	0.068	15.480	9.404	**2.177**	52.624	38.525	**3.826**
	49	0.114	0.047	**0.032**	15.378	8.976	**0.943**	51.837	36.880	**2.221**
	81	0.120	0.046	**0.019**	15.725	9.086	**0.544**	53.149	37.672	**1.486**

Furthermore, it is important to study the behaviour of AM-estimators for different sample size (N) and contamination (ϵ). Figures 5, 6 and 7 show the AM-estimators for $\alpha = \{-1, -6, -10\}$, considering $N = \{9, 25, 49, 81\}$ and $\epsilon = \{0.00, 0.01, 0.05, 0.10\}$, for an outlier of magnitude 40. In these figures, we do not observe significant differences between estimates based on small and large samples, allowing us to conclude that AM-estimators are very efficient in a wide range of situations.

Table 5. Numerical comparison of the mean among ML, M and AM estimators, for varying α, sample size and contamination level ϵ, with $\upsilon = 5$ (winning cases in boldface).

		$\alpha = -1$			$\alpha = -6$			$\alpha = -10$		
ϵ	N	$\hat{B}[\hat{\alpha}_{\mathrm{ML}}]$	$\hat{B}[\hat{\alpha}_{\mathrm{M}}]$	$\hat{B}[\hat{\alpha}_{\mathrm{AM}}]$	$\hat{B}[\hat{\alpha}_{\mathrm{ML}}]$	$\hat{B}[\hat{\alpha}_{\mathrm{M}}]$	$\hat{B}[\hat{\alpha}_{\mathrm{AM}}]$	$\hat{B}[\hat{\alpha}_{\mathrm{ML}}]$	$\hat{B}[\hat{\alpha}_{\mathrm{M}}]$	$\hat{B}[\hat{\alpha}_{\mathrm{AM}}]$
0.00	9	0.162	**0.140**	0.154	0.087	**0.086**	**0.086**	**0.008**	**0.008**	0.010
	25	0.048	**0.041**	0.045	**0.047**	**0.047**	**0.047**	0.031	0.031	**0.030**
	49	0.013	0.004	**0.002**	0.020	0.020	**0.017**	**0.017**	**0.017**	**0.017**
	81	0.014	**0.012**	0.013	**0.011**	**0.011**	0.012	**0.011**	**0.011**	**0.011**
0.01	9	0.174	0.080	**0.018**	**0.476**	0.478	0.481	**0.535**	0.538	0.383
	25	0.089	0.057	**0.042**	**0.250**	**0.250**	0.251	0.302	0.302	**0.163**
	49	0.060	0.043	**0.036**	**0.168**	**0.168**	**0.168**	0.192	0.192	**0.094**
	81	0.044	0.033	**0.029**	**0.122**	0.123	0.123	0.140	0.140	**0.064**
0.05	9	0.185	0.091	**0.035**	**0.507**	0.510	0.513	0.562	0.562	**0.430**
	25	0.143	0.100	**0.082**	**0.335**	0.336	0.337	0.387	0.387	**0.245**
	49	0.128	0.095	**0.086**	**0.284**	**0.284**	0.285	0.332	0.332	**0.201**
	81	0.125	0.092	**0.085**	**0.276**	0.277	0.278	0.329	0.329	**0.199**
0.10	9	0.208	0.111	**0.050**	**0.545**	0.549	0.553	0.599	0.600	**0.495**
	25	0.200	0.152	**0.137**	**0.439**	0.441	0.443	0.490	0.491	**0.363**
	49	0.216	0.173	**0.168**	**0.426**	0.427	0.429	0.484	0.485	**0.353**
	81	0.226	0.185	**0.183**	**0.440**	0.441	0.443	0.490	0.490	**0.355**

Table 6. Numerical comparison of the absolute relative bias among ML, M and AM estimators, for varying α, sample size and contamination level ϵ, with $\upsilon = 15$ (winning cases in boldface).

		$\alpha = -1$			$\alpha = -6$			$\alpha = -10$		
ϵ	N	$\hat{B}[\hat{\alpha}_{\mathrm{ML}}]$	$\hat{B}[\hat{\alpha}_{\mathrm{M}}]$	$\hat{B}[\hat{\alpha}_{\mathrm{AM}}]$	$\hat{B}[\hat{\alpha}_{\mathrm{ML}}]$	$\hat{B}[\hat{\alpha}_{\mathrm{M}}]$	$\hat{B}[\hat{\alpha}_{\mathrm{AM}}]$	$\hat{B}[\hat{\alpha}_{\mathrm{ML}}]$	$\hat{B}[\hat{\alpha}_{\mathrm{M}}]$	$\hat{B}[\hat{\alpha}_{\mathrm{AM}}]$
0.00	9	0.162	**0.140**	0.154	0.085	0.085	**0.084**	**0.000**	**0.000**	0.001
	25	0.048	**0.041**	0.045	**0.044**	**0.044**	**0.044**	0.030	0.030	**0.029**
	49	0.013	0.004	**0.002**	**0.019**	**0.019**	**0.019**	0.018	0.018	**0.017**
	81	0.014	**0.012**	0.013	**0.010**	**0.010**	**0.010**	**0.012**	**0.012**	**0.012**
0.01	9	0.318	**0.080**	0.258	0.697	0.533	**0.110**	0.770	0.666	**0.016**
	25	0.167	0.057	**0.044**	0.459	0.274	**0.049**	0.557	0.404	**0.021**
	49	0.106	0.043	**0.013**	0.326	0.177	**0.021**	0.404	0.262	**0.024**
	81	0.078	0.033	**0.006**	0.256	0.135	**0.008**	0.319	0.197	**0.019**
0.05	9	0.333	**0.091**	0.285	0.718	0.568	**0.101**	0.787	0.692	**0.006**
	25	0.233	0.100	**0.052**	0.550	0.369	**0.062**	0.631	0.492	**0.021**
	49	0.204	0.095	**0.022**	0.481	0.309	**0.027**	0.571	0.430	**0.023**
	81	0.198	0.092	**0.016**	0.474	0.303	**0.018**	0.565	0.423	**0.015**
0.10	9	0.362	**0.114**	0.354	0.741	0.606	**0.109**	0.804	0.720	**0.005**
	25	0.299	0.139	**0.085**	0.642	0.482	**0.055**	0.717	0.603	**0.007**
	49	0.319	0.170	**0.034**	0.644	0.482	**0.021**	0.712	0.593	**0.020**
	81	0.334	0.186	**0.019**	0.655	0.491	**0.012**	0.725	0.606	**0.015**

5. Conclusions and future work

In this work, we have introduced the AM-estimator, a novel robust method based on asymmetric robustifying functions. A Monte Carlo study was performed to investigate its robustness properties and to compare its performance with respect to maximum likelihood estimator and classical M-estimators. We estimate the roughness parameter α of the $\mathcal{G}_A^0$ distribution, and the results show that $\hat{\alpha}_{\mathrm{AM}}$ estimator is very favourable with respect to $\hat{\alpha}_{\mathrm{ML}}$ and $\hat{\alpha}_{\mathrm{M}}$ estimators under the mean square error criterion. Besides, we can see that $\hat{\alpha}_{\mathrm{AM}}$ has a smaller bias in the presence of contamination.

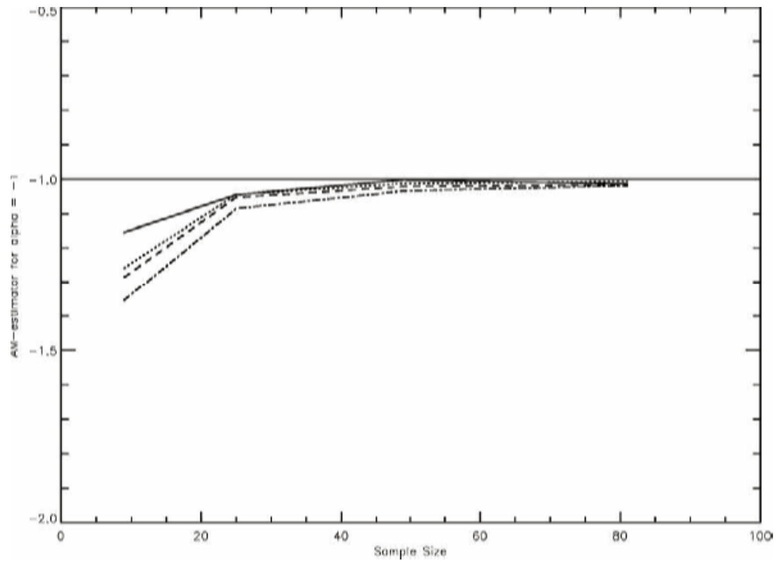


Figure 5. AM-estimators for $\alpha = -1$ versus sample size, varying contamination: $\epsilon = 0.00$ (solid line), $\epsilon = 0.01$ (dashed line), $\epsilon = 0.05$ (long-dashed line) and $\epsilon = 0.10$ (dashed long-dashed line).

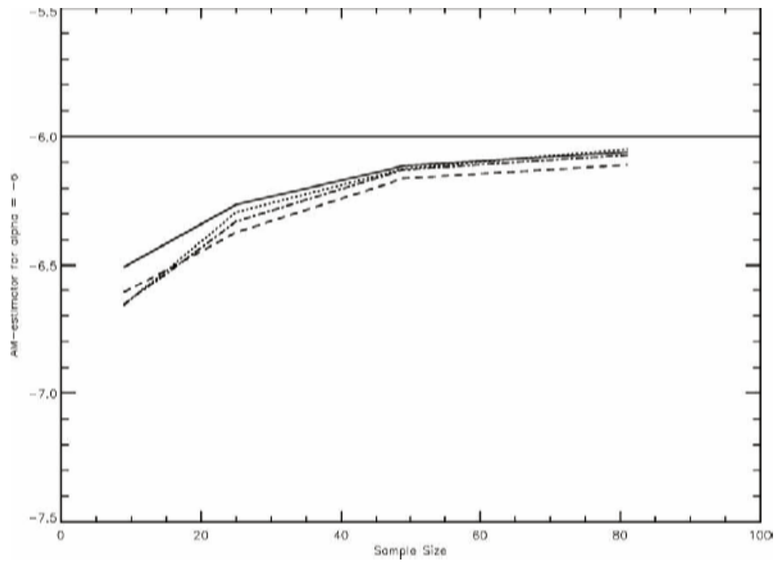


Figure 6. AM-estimators for $\alpha = -6$ versus sample size, varying contamination: $\epsilon = 0.00$ (solid line), $\epsilon = 0.01$ (dashed line), $\epsilon = 0.05$ (long-dashed line) and $\epsilon = 0.10$ (dashed long-dashed line).

The computational effort required to compute $\hat{\alpha}_{\mathrm{AM}}$ is comparable with the $\hat{\alpha}_{\mathrm{M}}$ estimator. Furthermore, it is noteworthy that the choice of the initial value is not critical for AM and M-estimators. The computation of the AM-estimators presents some numerical problems for small samples.

For the tuning parameters, there are several pairs of values of $r = (r_1, r_2)$ with the same relative asymptotic efficiency, and the rule does not determine r uniquely. As a remedy, we minimise the approximation of the maximum asymptotic variance as a function of r.

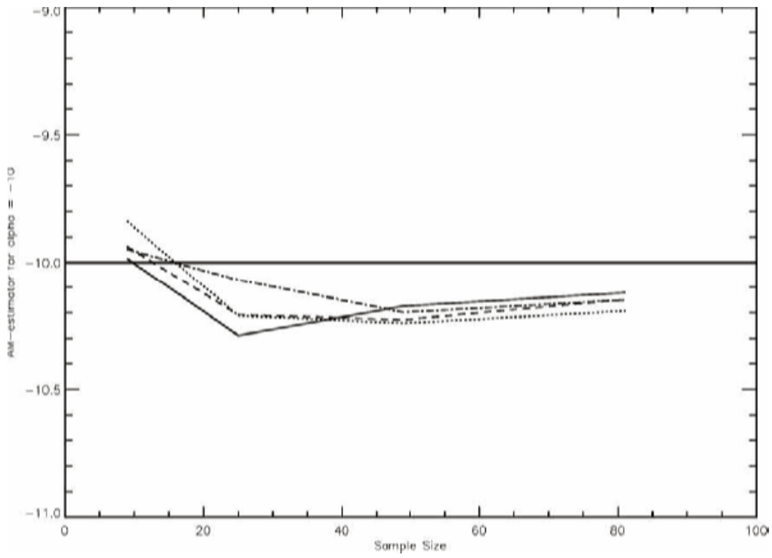


Figure 7. AM-estimators for $\alpha = -10$ versus sample size, varying contamination: $\epsilon = 0.00$ (solid line), $\epsilon = 0.01$ (dashed line), $\epsilon = 0.05$ (long-dashed line) and $\epsilon = 0.10$ (dashed long-dashed line).

This work will continue computing the AM-estimator for the multilook case, *i.e.* for the $L > 1$ situation, and for polarimetric multivariate data. Also, the simultaneous estimation of the α and the γ parameters will be considered. Another future work is to study estimators for confidence intervals, for instance, using bootstrap or other resampling methods.

Acknowledgements

This work was partially supported by Fondecyt Research Grant No. 1010101, by Universidad Técnica Federico Santa María DGIP Grant No. 240224 and by CNPq.

References

[1] Allende, H., Galbiati, J. and Vallejos, R., 2001, Robust image modelling on image processing. *Pattern Recognition Letters*, **22**(11), 1219–1231.
[2] Oliver, C. and Quegan, S., 1998, *Understanding Synthetic Aperture Radar Images* (Boston: Artech House).
[3] Frery, A.C., Müller, H.-J., Yanasse, C.C.F. and Sant'Anna, S.J.S., 1997, A model for extremely heterogeneous clutter. *IEEE Transactions on Geoscience and Remote Sensing*, **35**(3), 648–659.
[4] Mejail, M.E., Frery, A.C., Jacobo-Berlles, J. and Bustos, O.H., 2001, Approximation of distributions for SAR images: proposal, evaluation and practical consequences. *Latin American Applied Research*, **31**, 83–92.
[5] Navarrete, H.A., Frery, A.C., Sánchez, F. and Antó, J., 2002, Ultrasound images filtering using the multiplicative model. In: M.F. Insana and W.F. Walker (Eds) *Progress in Biomedical Optics and Imaging* (San Diego, USA: SPIE), pp. 345–353. Proceedings of SPIE vol. 4687.
[6] Frery, A.C., Sant'Anna, S.J.S., Mascarenhas, N.D.A. and Bustos, O.H., 1997, Robust inference techniques for speckle noise reduction in 1-look amplitude SAR images. *Applied Signal Processing*, **4**, 61–76.
[7] Jin, J.Y., Silva, G.T. and Frery, A.C., 2004, SAR despeckling filters in ultrasound imaging. *Latin American Applied Research*, **34**, 49–53. Also available online at http://www.ufal.br/sites/tci/members/frery/publications/journalpapers/JinSilvaFreryUltrasoundLAAR.
[8] Lucini, M.M., Ruiz, V.F., Frery, A.C. and Bustos, O.H., 2003, Robust classification of SAR imagery. IEEE International Conference on Acoustics, Speech and Signal Processing, IEEE, Hong Kong, pp. 557–560.
[9] Mejail, M.E., Jacobo-Berlles, J., Frery, A.C. and Bustos, O.H., 2003, Classification of SAR images using a general and tractable multiplicative model. *International Journal of Remote Sensing*, **24**, 3565–3582.
[10] Polzehl, J. and Spokoiny, V., 2003, Image denoising: ponitwise adaptive approach. *The Annals of Statistics*, **31**(1), 30–57.
[11] Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J. and Stahel, W.A., 1986, *Robust Statistics: the Approach Based on Influence Functions* (New York: Wiley).

[12] Nikias, C.L. and Shao, M., 1995, *Signal Processing with Alpha-Stable Distributions and Applications* (Wiley).
[13] Reed III, J.F. and Stark, D.B., 1996, Hinge estimators of location: robust to asymmetry. *Computer Methods and Programs in Biomedicine*, **49**, 11–17.
[14] Cribari-Neto, F., Frery, A.C. and Silva, M.F., 2002, Improved estimation of clutter properties in speckled imagery. *Computational Statistics and Data Analysis*, **40**(4), 801–824.
[15] Bustos, O.H., Lucini, M.M. and Frery, A.C., 2002, M-estimators of roughness and scale for GA0-modelled SAR imagery. *EURASIP Journal on Applied Signal Processing*, **2002**(1), 105–114.
[16] Joughin, I.R., Percival, D.B. and Winebrenner, D.P., 1993, Maximum likelihood estimation of K distribution parameters for SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, **31**(5), 989–999.
[17] Huber, P.J., 1981, *Robust Statistics* (New York: Wiley).
[18] Andrews, D., Bickel, P., Hampel, F., Huber, P.J., Rogers, W. and Tuckey, J., 1972, *Robust Estimates of Location: Survey and Advances* (New Jersey: Princeton University Press).
[19] Rousseeuw, P.J. and Verboven, S., 2002, Robust estimation in very small samples. *Computational Statistics and Data Analysis*, **40**(4), 741–758.
[20] Marazzi, A. and Ruffieux, C., 1996, *Implementing M-Estimators of the Gamma Distribution* (Berlin: Springer-Verlag) Lecture Notes in Statistics, 109.