

# M-Estimators of Roughness and Scale for $G_A^0$ -Modelled SAR Imagery

**Oscar H. Bustos**

*Facultad de Matemática, Astronomía y Física, Universidad Nacional de Córdoba, Ciudad Universitaria, 5000 Córdoba, Argentina*  
Email: bustos@mate.uncor.edu

**María Magdalena Lucini**

*Facultad de Matemática, Astronomía y Física, Universidad Nacional de Córdoba, Ciudad Universitaria, 5000 Córdoba, Argentina*  
Email: lucini@mate.uncor.edu

**Alejandro C. Frery**

*Centro de Informática, Universidade Federal de Pernambuco, CP 7851, 50732-970 Recife - PE, Brasil*  
Email: frery@cin.ufpe.br

Received 31 July 2001 and in revised form 10 October 2001

The  $G_A^0$  distribution is assumed as the universal model for multilook amplitude SAR imagery data under the multiplicative model. This distribution has two unknown parameters related to the roughness and the scale of the signal, that can be used in image analysis and processing. It can be seen that maximum likelihood and moment estimators for its parameters can be influenced by small percentages of “outliers”; hence, it is of outmost importance to find robust estimators for these parameters. One of the best-known classes of robust techniques is that of M-estimators, which are an extension of the maximum likelihood estimation method. In this work we derive the M-estimators for the parameters of the  $G_A^0$  distribution, and compare them with maximum likelihood estimators with a Monte-Carlo experience. It is checked that this robust technique is superior to the classical approach under the presence of corner reflectors, a common source of contamination in SAR images. Numerical issues are addressed, and a practical example is provided.

**Keywords and phrases:** inference, likelihood, M-estimators, Monte-Carlo method, multiplicative model, speckle, synthetic aperture radar, robustness.

## 1. INTRODUCTION

The statistical modeling of synthetic aperture radar (SAR) imagery has provided some of the best tools for the processing and understanding of this kind of data. Among the statistical approaches the most successful is the multiplicative model.

This model offers a set of distributions that, with a few parameters, are able to characterize most of SAR data. This model is presented, for instance, in [1], and extended in [2].

This extension is a general and tractable set of distributions within the multiplicative model, used to describe every kind of SAR return. It was then called a universal model, and its properties are studied in [3, 4, 5].

In this paper, the problem of estimating the parameters of this extension, namely of the  $G_A^0$  distribution, is studied for the single look (the noisiest) situation. Two typical estimation situations arise in image processing and analysis, namely large and small samples, being the latter considered in this work. The small samples problem arises in, for instance, im-

age filtering where with a few observations within a window a new value is computed.

Statistical inference with small samples is subjected to many problems, mainly bias, large variance, and sensitivity to deviations from the hypothesized model. This last issue is also a problem when dealing with large samples.

Robustness is a quite desirable property for estimators, since it allows their use even in situations where the quality of the input data is not perfect [6, 7, 8]. Most image processing and analysis procedures (filtering, classification, segmentation, etc.) use spatial data. Even experienced users are unable to guarantee that all the input data are free of spurious values and/or structures.

A situation where this occurs is in the presence of corner reflectors. These devices, which are essential for data calibration, produce a return which is quite higher than the rest of the image. If data from a corner reflector enter a non-robust estimation procedure, the results may be completely

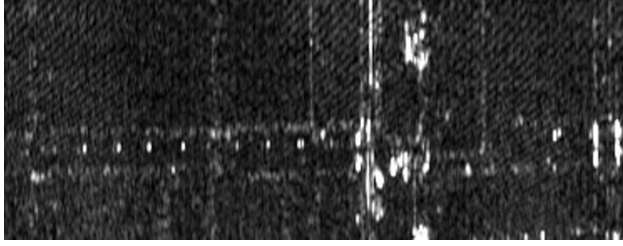


FIGURE 1: Man-made corner reflectors on a SAR image.

unreliable. Figure 1 shows man-made corner reflectors, the white spot regularly placed horizontally along the middle of the image, possibly buildings, on a SAR image where crops predominate.

This paper presents classical estimators, namely those based on sample moments and the maximum likelihood ones, and derives robust M-estimators for the single look  $\mathcal{G}_A^0$  model. Once these estimators are derived, several situations are compared by means of a Monte-Carlo experience. These estimators are then applied to SAR imagery, and it is shown that the robust approach exhibits a good performance even in the presence of corner reflectors.

M-estimators have been mainly used for symmetric data, being [7, 8] two very relevant exceptions. In this application they are computed, implemented, and assessed for speckled imagery that, as will be seen in the next section, follow laws whose densities can be highly asymmetric. Since speckle noise also appears in ultrasound B-scan, sonar and laser images, the procedures here presented have potential application in all these techniques.

## 2. NOTATION AND MODEL DEFINITION

In the following,  $1_A$  will denote the indicator function of the set  $A$ , that is,

$$1_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{else.} \end{cases} \quad (1)$$

The single-look  $\mathcal{G}_A^0(\alpha, \gamma)$  distribution is characterized by the following probability density function:

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

Multilook, intensity, and complex versions of this distribution can be seen in [2]. The polarimetric (multivariate) extension of the  $\mathcal{G}^0$  distribution is presented in [9].

The parameter  $\alpha$  in (2) describes the roughness of the target, being small values (say  $\alpha < -15$ ) usually associated to homogeneous targets, like pasture, values ranging in the  $[-15, -5]$  interval usually observed in heterogeneous clutter, like forests, and big values ( $-5 < \alpha < 0$  for instance) commonly seen when extremely heterogeneous areas, as urban spots, are imaged. The  $\gamma$  parameter is related to the scale, in the sense that if  $Z'_A$  is a  $\mathcal{G}_A^0(\alpha, 1)$ -distributed random vari-

able then  $Z = \sqrt{\gamma}Z'$  obeys a  $\mathcal{G}_A^0(\alpha, \gamma)$  law and, therefore, it is related to the brightness of the scene. Both parameters are essential when characterizing targets, and in image processing techniques involving statistical modeling.

In the following it will be assumed that  $\alpha < 0$  and  $\gamma > 0$ . Calling

$$\begin{aligned} s_1(x, (\alpha, \gamma)) &= \frac{\partial}{\partial \alpha} \ln f(x, (\alpha, \gamma)), \\ s_2(x, (\alpha, \gamma)) &= \frac{\partial}{\partial \gamma} \ln f(x, (\alpha, \gamma)), \end{aligned} \quad (3)$$

the score function of the  $\mathcal{G}_A^0(\alpha, \gamma)$  distribution is given by

$$\begin{aligned} \mathbf{s}(x, (\alpha, \gamma)) &= \begin{pmatrix} s_1(x, (\alpha, \gamma)) \\ s_2(x, (\alpha, \gamma)) \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{\alpha} + \ln \left( 1 + \left( \frac{x}{\sqrt{\gamma}} \right)^2 \right) \\ -\frac{\alpha}{\gamma} - \frac{1-\alpha}{\gamma + x^2} \end{pmatrix}. \end{aligned} \quad (4)$$

The cumulative distribution function of such random variable is given by

$$F(x, (\alpha, \gamma)) = \left( 1 - \left( 1 + \frac{x^2}{\gamma} \right)^\alpha \right) 1_{(0,+\infty)}(x). \quad (5)$$

This will be used to compute an estimator based on order statistics. In [5] a relation between a more general version of the  $\mathcal{G}_A^0(\alpha, \gamma)$  law and Snedekor's  $\mathcal{F}$  distribution is shown to allow writing (5) as a function of the cumulative distribution function of the latter.

The moments of a  $\mathcal{G}_A^0(\alpha, \gamma)$  distribution are given by

$$E(X^k) = \begin{cases} \gamma^{k/2} \frac{k\Gamma(k/2)\Gamma(-\alpha - k/2)}{2\Gamma(-\alpha)} & \text{if } -\alpha > \frac{k}{2}, \\ \infty & \text{else.} \end{cases} \quad (6)$$

This distribution can be derived as the square root of the ratio of two independent random variables: one obeying a unitary-mean exponential distribution (which conveys the speckle noise in one look) and one following a  $\Gamma(\alpha, \gamma)$  distribution, related to the unobserved ground truth, the backscatter. Densities of the  $\mathcal{G}_A^0(\alpha, \gamma)$  distribution are shown in Figure 2, for  $\alpha \in \{-1, -3, -9\}$  and the scale parameter  $\gamma = \gamma_\alpha$  adjusted to have unitary mean, that is, using (6) and  $k = 1$ , one derives

$$\gamma_\alpha = \frac{4}{\pi} \left( \frac{\Gamma(-\alpha)}{\Gamma(-\alpha - 1/2)} \right)^2. \quad (7)$$

Following Barndorff-Nielsen and Blæsild [10], it is interesting to see these densities as log probability functions, particularly because the  $\mathcal{G}_A^0$  law is closely related to the class of hyperbolic distributions [11]. Figure 3 shows the densities of the  $\mathcal{G}_A^0(-2.5, 7.0686/\pi)$  and the Gaussian distribution  $\mathcal{N}(1, 4(1.1781 - \pi/4)/\pi)$  distributions in semilogarithmic scale, along with their mean value. The parameters were

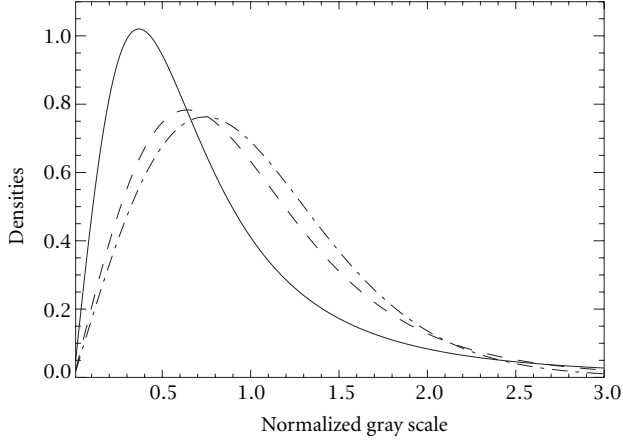


FIGURE 2: Densities of the  $\mathcal{G}_A^0(\alpha, \gamma_\alpha)$  distribution, with  $\alpha \in \{-1, -3, -9\}$  (solid line, dashes, dash-dot, respectively) with unitary mean.

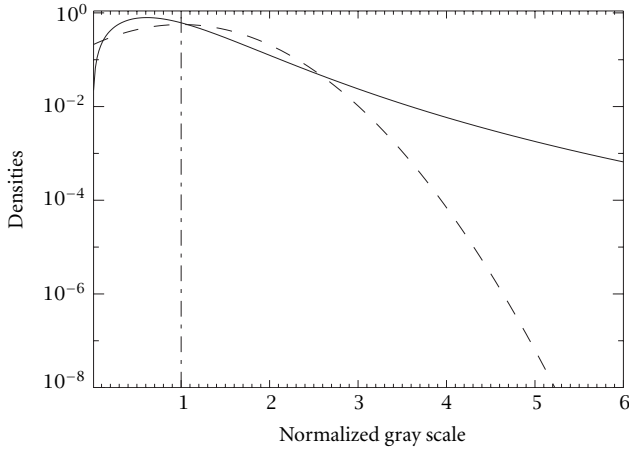


FIGURE 3: Densities of the  $\mathcal{G}_A^0(-2.5, 7.0686/\pi)$  (solid line) and the  $\mathcal{N}(1, 4(1.1781 - \pi/4)/\pi)$  (dashes) distributions in semilogarithmic scale, with mean value in dash-dot.

chosen, using (6), so that these distributions have equal mean and variance. The different decays of their tails in the logarithmic plot are evident: the former behaves logarithmically, while the latter decays quadratically. This behavior ensures the ability of the  $\mathcal{G}_A^0$  distribution to model data with extreme variability.

### 3. INFERENCE FOR THE $\mathcal{G}_A^0$ MODEL

It can be seen that the maximum likelihood estimators of  $\alpha$  and  $\gamma$  based in  $X_1, \dots, X_N$ , represented here by  $\hat{\alpha}_{ML,N}$  and  $\hat{\gamma}_{ML,N}$ , respectively, are given by

$$\hat{\alpha}_{ML,N} = - \left( \frac{1}{N} \sum_{k=1}^N \ln \left( 1 + \left( \frac{X_k}{\sqrt{\hat{\gamma}_{ML,N}}} \right)^2 \right) \right)^{-1} \quad (8)$$

and by

$$\hat{\gamma}_{ML,N} = \left( 1 + 2 \left( \frac{1}{N} \sum_{k=1}^N \ln \left( 1 + \left( \frac{X_k}{\sqrt{\hat{\gamma}_{ML,N}}} \right)^2 \right) \right)^{-1} \right) \times \frac{1}{N} \sum_{k=1}^N \frac{X_k^2}{1 + (X_k/\sqrt{\hat{\gamma}_{ML,N}})^2}. \quad (9)$$

Assume  $\alpha < -1/2$  in order to have random variables with finite mean. For each  $j = 1/2, 1$ , define the  $j$ th order moment as  $m_{j,N} = N^{-1} \sum_{k=1}^N X_k^j$ . Using (6) it can be seen that the half and first order moments estimators of  $\alpha$  and  $\gamma$  based on  $X_1, \dots, X_N$ , denoted here by  $\hat{\alpha}_{MOM,N}$  and by  $\hat{\gamma}_{MOM,N}$ , respectively, are given by the solution of

$$m_{1/2,N} = \hat{\gamma}_{MOM,N}^{1/4} \frac{\Gamma(-\hat{\alpha}_{MOM,N} - 1/4)\Gamma(1/4)}{4\Gamma(-\hat{\alpha}_{MOM,N})}, \quad (10)$$

$$m_{1,N} = \hat{\gamma}_{MOM,N}^{1/2} \frac{\sqrt{\pi}\Gamma(-\hat{\alpha}_{MOM,N} - 1/2)}{2\Gamma(-\hat{\alpha}_{MOM,N})}.$$

Moment estimation is extensively used in remote sensing applications [2], mainly because it is inexpensive from the computational point of view and analytically tractable in most situations. In the presence of corner reflectors, though, severe numerical instabilities were observed.

A comparison among different estimators for roughness parameter of the  $\mathcal{G}_A^0$  distribution was performed in [4] through a Monte-Carlo experience. No contamination was considered, and the  $\hat{\alpha}_{ML}$  estimator was the best one in almost all cases with respect to the mean square error (*mse*) and the closeness to the true value (*bias*) criteria, assuming  $\gamma$  is known. No analytical comparison is possible, due to the complexity of the estimators.

In previous works (see [12], for instance) it is shown that moment and maximum likelihood estimators have many optimal properties when the observations,  $x_1, \dots, x_N$  are the outcome from independent, identically distributed random variables, with common density  $f(\cdot, (\alpha, \gamma))$ . Among these properties, maximum likelihood estimators are asymptotically unbiased, that is, if  $X_1, \dots, X_N$  is a sequence of independent, identically distributed random variables with common density  $f(\cdot, (\alpha, \gamma))$  then, when  $N \rightarrow \infty$ , it holds that  $\hat{\alpha}_{MOM,N} \rightarrow \alpha$ ,  $\hat{\gamma}_{MOM,N} \rightarrow \gamma$ ,  $\hat{\alpha}_{ML,N} \rightarrow \alpha$ , and  $\hat{\gamma}_{ML,N} \rightarrow \gamma$ . Nevertheless, good performance is neither warranted with finite samples nor if the sample does not obey precisely the hypothesis. A common departure from classical hypothesis is the contamination by a percentage of “outliers,” that is, when some of the observed data come from a different distribution.

“Corner reflectors” can be considered outliers in SAR imagery. These are physical artifacts in the sensed area that return most of the power they receive. The image in these areas is dominated by the biggest possible values admitted by the storage characteristics, and their effect is typically limited to a few pixels. Corner reflectors are either placed on purpose, for image calibration, due to man-made objects, such as highly reflective urban areas, or the result of double-bounce reflection [1].

Since the routines that compute both ML- and M-estimators require initial guesses, the unstable (and, therefore, unacceptable) behavior of moment estimators demanded the use of another technique. A procedure based on analogy using order statistics and moments [13], defined in the following for  $\alpha < -1$ , was used.

If  $X$  is  $\mathcal{G}_A^0(\alpha, \gamma)$  distributed the median of the scaled random variable  $Y = X/E(X)$  can be computed using (5)

$$Q_2 = \frac{2}{\sqrt{\pi}} \sqrt{\sqrt[3]{1/2} - 1} \frac{\Gamma(-\alpha)}{\Gamma(-\alpha - 1/2)}. \quad (11)$$

This scaled median does not depend on  $\gamma$ , so  $\hat{\alpha}$  can be estimated using the sample median  $q_2(\mathbf{y})$ , where  $\mathbf{y} = (x_i/\bar{\mathbf{x}})_i$ , using standard numerical tools. An estimate of  $\gamma$  using the first-order moment can then be computed as

$$\hat{\gamma} = \frac{4}{\pi} m_{1,N}^2 \left( \frac{\Gamma(-\hat{\alpha})}{\Gamma(-\hat{\alpha} - 1/2)} \right)^2. \quad (12)$$

This estimator of  $\alpha$  derived from (11) and the one computed for  $\gamma$  through (12) will be called *mixed estimator* for  $(\alpha, \gamma)$ , and denoted  $(\hat{\alpha}_{\text{MIX}}, \hat{\gamma}_{\text{MIX}})$ .

In the following, ML- and M-estimators will be assessed in two situations: the pure model, when no contamination is present, and cases where outliers simulating a corner reflector enter the data.

The contamination model here considered is defined as a sequence of independent, identically distributed random variables  $X_1, \dots, X_N$  with common distribution function  $F(\cdot, (\alpha, \gamma), \epsilon, z)$  given by

$$F(x, (\alpha, \gamma), \epsilon, z) = (1 - \epsilon)F(x, (\alpha, \gamma)) + \epsilon\delta_z(x), \quad (13)$$

where  $\delta_z(x) = 1_{[z, +\infty)}(x)$ , with  $z$  a “very big” value with respect to most of the values typically assumed by a random variable with distribution  $\mathcal{G}_A^0(\alpha, \gamma)$ . Equation (13) describes a contamination that occurs at random with probability  $\epsilon$ , that is, in average  $N\epsilon$  out of  $N$  samples will be “very big” values (corner reflectors), while  $N - N\epsilon$  samples will obey the  $\mathcal{G}_A^0(\alpha, \gamma)$  distribution (the background). The flexibility of the  $\mathcal{G}_A^0(\alpha, \gamma)$  distribution will allow the modeling of any kind of background, namely crops, forests, and urban areas. The contamination value  $z$  will be chosen as a factor of the mean value of the underlying distribution  $\mathcal{G}_A^0(\alpha, \gamma)$ .

Other contamination hypothesis may include different distributions for the departure of the model, and/or spatial dependence among observations.

In order to quantitatively assess the sensitivity of ML-estimators in our case, using the strong law of large numbers on (8) and assuming  $\gamma = 1$ , it is immediate that

$$\begin{aligned} \lim_{N \rightarrow \infty} \hat{\alpha}_{\text{ML},N}(z) &= \left( E_{F(\cdot, (\alpha, \gamma), \epsilon, z)}(\ln(1 + X^2)) \right)^{-1} \\ &= \frac{1}{\epsilon \ln(1 + z^2) - (1 - \epsilon)(1/\alpha)}. \end{aligned} \quad (14)$$

Figure 4 shows, for  $\alpha = -5$ , this limit as a function of  $\epsilon$ , the proportion of outliers in the sample, for several values of

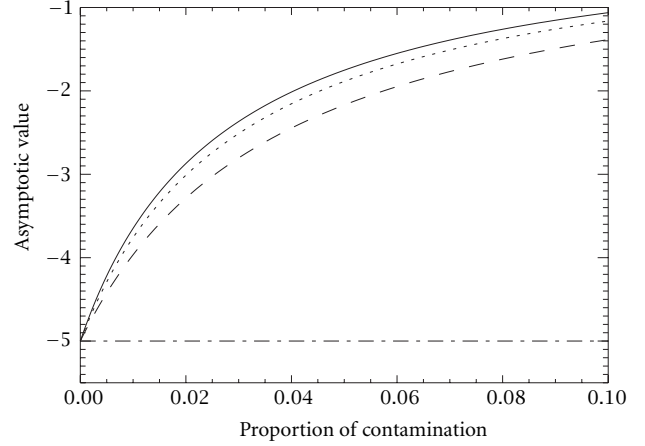


FIGURE 4: Asymptotic value of  $\hat{\alpha}_{\text{ML}}$ , under the contaminated model with  $\alpha = -5$ , as a function of the proportion of contamination  $\epsilon$  and outlier values  $-45$  (solid line),  $-30$  (dots), and  $-15$  (dashes).

the outlier  $z$ . Three values of  $z$  are shown for  $\epsilon \in [0, 0.1]$ . It can be seen that a sample that suffers from a small contamination of  $\epsilon = 0.02$  leads to the wrong conclusion that extremely heterogeneous targets are under observation, since the estimated value will be around  $-3$  whilst the true value is  $-5$  (corresponding to an heterogeneous target). The bigger the contamination the worse this behavior. This influence on the ML-estimator is noticed regardless the value of  $z$  in the chosen range. This result justifies a careful analysis of the performance of estimators for a single representative value of the contamination  $z$ , and for several values of  $\epsilon$ , as presented in Section 5.

#### 4. M-ESTIMATORS

These estimators offer a useful alternative when small proportions of values may be far from the bulk of data. A good survey on these estimators and their use in practical situations can be seen in [7]. The difficulty in our case arises due to the fact that the underlying distribution is asymmetric. Inspired in the work by Marazzi and Ruffieux [8], that robustified maximum likelihood estimators for the parameters of the gamma distribution with success, we will compute M-estimators for the parameters of the  $\mathcal{G}_A^0(\alpha, \gamma)$  distribution.

One has to devise ways to prevent a large influence of outliers in the likelihood equations. M-estimators propose the use of  $\psi$  functions that truncate the score of influential observations in the likelihood equations.

Let  $b_1$  and  $b_2$  be two real positive numbers. We will call M-estimators of parameters  $\alpha$  and  $\gamma$  based on the sample  $X_1, \dots, X_N$ , the statistics  $\hat{\alpha}_{\text{M},N}$ , and  $\hat{\gamma}_{\text{M},N}$  such that

$$\begin{aligned} \sum_{k=1}^N \psi_{b_1}(s_1(X_k, (\hat{\alpha}_{\text{M},N}, \hat{\gamma}_{\text{M},N})) - c_1(\hat{\alpha}_{\text{M},N}, \hat{\gamma}_{\text{M},N}, b_1)) &= 0, \\ \sum_{k=1}^N \psi_{b_2}(s_2(X_k, (\hat{\alpha}_{\text{M},N}, \hat{\gamma}_{\text{M},N})) - c_2(\hat{\alpha}_{\text{M},N}, \hat{\gamma}_{\text{M},N}, b_2)) &= 0, \end{aligned} \quad (15)$$

where  $s_1$  and  $s_2$  are given in (4), and  $\psi_{b_1}$  and  $\psi_{b_2}$  are

$$\psi_{b_i}(x) = \begin{cases} -b_i & \text{if } x \leq b_i, \\ x & \text{if } -b_i < x < b_i, \\ b_i & \text{if } x \geq b_i, \end{cases} \quad (16)$$

with  $x \in \mathbb{R}$  and  $i = 1, 2$ . Note that making these functions the identity and  $c_1 = c_2 \equiv 0$ , equations (15) reduce to the familiar system of likelihood equations. Figure 5 illustrates the  $\psi_6(x)$  function, where its effect on the data becomes clear: it truncates the score of those values above and below its threshold. In this manner, this function prevents abnormal (too small and too big) data having excessive influence on the estimates.

The functions  $c_i : \Theta \times (0, +\infty) \rightarrow \mathbb{R}$  in (15) are defined in such way that  $(\hat{\alpha}_{M,N}, \hat{\gamma}_{M,N})_N$  is a sequence of asymptotically unbiased estimators of  $(\alpha, \gamma)$ , that is,

$$\begin{aligned} \int \psi_{b_1}(s_1(x, (\alpha, \gamma)) - c_1(\alpha, \gamma, b_1))f(x, (\alpha, \gamma))dx &= 0, \\ \int \psi_{b_2}(s_2(x, (\alpha, \gamma)) - c_2(\alpha, \gamma, b_2))f(x, (\alpha, \gamma))dx &= 0, \end{aligned} \quad (17)$$

for any  $(\alpha, \gamma) \in \Theta$ ,  $b_1 > 0$  and  $b_2 > 0$ . The computation of these functions is presented in Appendices A and B.

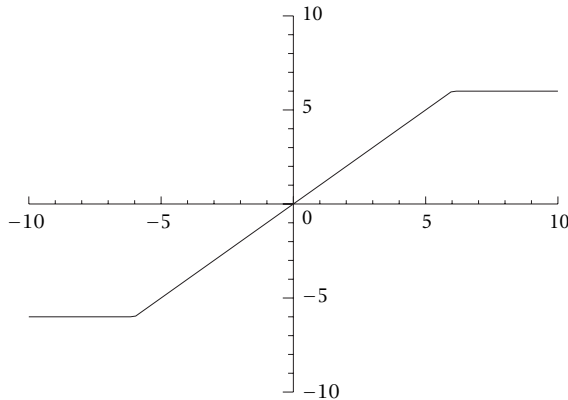


FIGURE 5: The  $\psi_6$  function used to define M-estimators.

The values  $b_1$  and  $b_2$ , called “tuning parameters” are chosen in such a way that the efficiency of the M-estimators is not unacceptably poor with respect to the maximum likelihood ones. Thus, we will choose  $b_1$  and  $b_2$  such that

$$\frac{V(\hat{\alpha}_{ML})}{V(\hat{\alpha}_M)} \cong e_1, \quad \frac{V(\hat{\gamma}_{ML})}{V(\hat{\gamma}_M)} \cong e_2, \quad (18)$$

where  $V$  denotes variance, with  $0.9 \leq e_i \leq 0.975$  for example. This restriction imposes that the variance of the robust estimators will not surpass those of the maximum likelihood ones in more than a certain factor. The computation of these parameters is done in Appendix C.

In order to assess the finite sample behavior of the proposed estimators consider the situation where  $N$  observations from independent identically distributed  $G_A^0(\alpha, \gamma)$  random

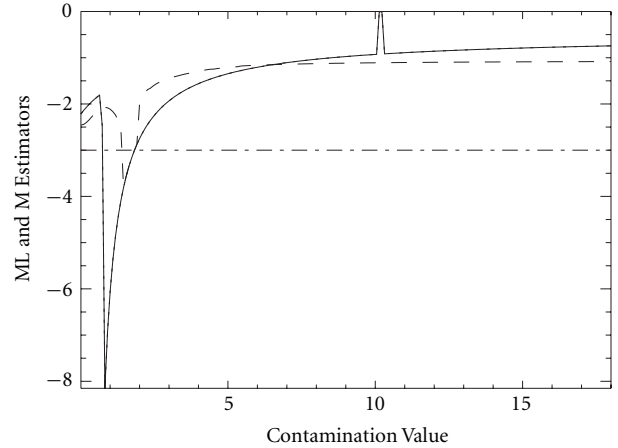


FIGURE 6: Maximum likelihood (continuous line) and M-(dotted line) estimators for the estimation of  $\alpha$  and  $\gamma$  in a sample with  $N = 77$  pure observations and  $M = 4$  contaminated values.

variables are contaminated by  $L$  observations taking the value  $z$ . The contaminated sample  $\mathbf{z} = (z_1, \dots, z_N, \underbrace{z, \dots, z}_L)$ , where the brackets denote the  $N + L$  outliers, will be used to estimate  $\alpha$  and  $\gamma$ . Depending on the true parameters, on the contamination (the value  $z$  and the number of outliers  $L$ ) and on the sample size  $N$ , it is expected that any non-robust estimator will be misled by the departure from the hypothesized model. Figure 6 shows the maximum likelihood and M-estimators for the situation where  $N = 77$  pure observations and  $L = 4$  equal contaminated values are used. It can be seen that the former suffers from both the departure from the true model and from numerical instabilities (the peak around the value 10). It is noticeable that the robust estimator remains closer to the true value (the dash-dot line) than the other in the presence of contamination. The bigger the value of the contamination the further maximum likelihood will go from the true value, while the M-estimator will render the same estimate. The same behavior was observed for the estimators of  $\gamma$ . The values employed for  $N$ ,  $L$ ,  $\alpha$  and  $z$  are representative of the kind of problem at hand: filtering images over a heterogeneous area ( $\alpha = -3$ ) with  $9 \times 9$  windows ( $N + L = 81$ ) where some observations ( $L = 4$ ) are due to a corner reflector ( $z$  varying and taking large values).

## 5. RESULTS

A Monte-Carlo experience was performed to compare maximum likelihood (ML) and M-estimators (M), since analytical comparisons are unfeasible.  $G_A^0$  deviates can be obtained in two ways, namely (a) by the inversion of the cumulative distribution function, given in (5), or (b) by the constructive definition of this distribution, that is, returning  $z = \sqrt{\gamma/x}$  where  $\gamma$  is a sample of the unitary mean exponential distribution and  $x$  is a sample of the  $\Gamma(-\alpha, \gamma)$  distribution,  $X$  independent of  $Y$ . Alternative (b) was chosen in this experience. Since the parameter  $\gamma$  is a scale factor, all the studied situations will use the parameter that makes the expected

value 1, that is,  $y = y_\alpha = 4(\Gamma(-\alpha)/\Gamma(-\alpha - 1/2))^2/\pi$ .

We studied four models that will be identified as Models 1, 2, 3, and 4: Model 1 (the “pure” model) consisting of samples of the  $F(\cdot, (\alpha, \gamma_\alpha))$  distribution; Model 2 consisting of samples taken from the  $F(\cdot, (\alpha, \gamma_\alpha), \epsilon, z)$  distribution with  $\epsilon = 0.01$  and  $z = 15$ , Model 3 is a collection of samples obtained from the  $F(\cdot, (\alpha, \gamma_\alpha), \epsilon, z)$  distribution with  $\epsilon = 0.05$  and  $z = 15$ , while Model 4 is a collection of samples obtained from the  $F(\cdot, (\alpha, \gamma_\alpha), \epsilon, z)$  distribution with  $\epsilon = 0.10$  and  $z = 15$ . In this manner, the pure situation and three levels of contamination are examined. The contamination value  $z$  is held fixed, as discussed in Section 3, as fifteen times the expected value of the underlying model which is big enough to describe corner reflectors.

In each of these models we deal with three situations, namely with: Situation 1 ( $\alpha = -1.0$ ), Situation 2 ( $\alpha = -3$ ), and Situation 3 ( $\alpha = -6$ ). These situations cover representative targets in practice: extreme heterogeneity, heterogeneity and homogeneity.

In order to determine the number of replications an empirical precision criterion was used. One hundred replications ( $1 \leq r \leq R' = 100$ ) with samples of size  $N = 9 \times 9 = 81$  were performed. One estimator for  $\alpha$  is computed for each replication  $r$ , say  $\hat{\alpha}(r)$ . The final number of replications  $R = R' + \ell$ ,  $\ell \geq 1$  integer, is the smallest number of samples that allows forming an asymptotic confidence interval at the 95% level on the mean of the  $R$  estimators, that is,  $2 \times 1.96s_R/\sqrt{R} < 0.5$  where  $s_R$  is the standard deviation of the sample  $\{\hat{\alpha}(1), \dots, \hat{\alpha}(R)\}$ . This is computed for every estimator for  $\alpha$  and every estimator for  $\gamma$  in every model and every situation. The biggest number of replications  $R$  (the worst case) is then used. This procedure led to  $R = 101$  in every model and every situation.

The following tables show the results obtained in this study for each situation: for every model the sample mean of the estimator over the  $R$  replications  $\hat{E}(\zeta) = R^{-1} \sum_{r=1}^R \zeta(r)$ , the mean square error  $mse = \hat{E}(\zeta - \theta)^2 = \hat{V}(\zeta) + (\hat{E}(\zeta) - \theta)^2$ , and the absolute relative bias  $B(\zeta) = \theta^{-1} |\hat{E}(\zeta) - \theta|$ , where  $\theta$  is the true value of the parameter and  $\zeta$  is the estimator under study.

In all the tables it can be seen that M-estimators are almost as good as ML-estimators when there is no contamination, while in the presence of outliers M-estimators exhibit smaller absolute relative bias and smaller mean squared error. As  $\alpha$  is smaller (the observed region is more homogeneous), estimation becomes less reliable. This is probably due to the shape of the likelihood function, that becomes flat and, therefore, is hard to find its maximum location. The bigger the proportion of contamination the worse the behavior of both estimators, but M-estimators remain consistently closer to the true value than ML-estimators, as expected.

TABLE 2: Model 2 ( $\epsilon = 0.01$ ) and three situations.

		$\hat{E}(\hat{\alpha})$	$B(\hat{\alpha})$	$mse(\hat{\alpha})$	$\hat{E}(\hat{\gamma})$	$B(\hat{\gamma})$	$mse(\hat{\gamma})$
$\alpha = -1$	ML	-0.97	0.03	0.22	0.41	0.01	0.08
	M	-1.01	0.01	0.15	0.49	0.20	0.08
$\alpha = -3$	ML	-2.96	0.01	25.07	2.72	0.05	46.06
	M	-3.10	0.03	3.28	3.17	0.10	5.65
$\alpha = -6$	ML	-2.76	0.54	16.10	2.89	0.57	22.32
	M	-4.28	0.28	7.41	4.77	0.29	9.43

TABLE 3: Model 3 ( $\epsilon = 0.05$ ) and three situations.

		$\hat{E}(\hat{\alpha})$	$B(\hat{\alpha})$	$mse(\hat{\alpha})$	$\hat{E}(\hat{\gamma})$	$B(\hat{\gamma})$	$mse(\hat{\gamma})$
$\alpha = -1$	ML	-0.68	0.32	0.12	0.27	0.33	0.03
	M	-0.86	0.14	0.12	0.39	0.04	0.08
$\alpha = -3$	ML	-0.89	0.70	4.45	0.70	0.76	4.87
	M	-2.01	0.33	1.46	2.13	0.26	1.60
$\alpha = -6$	ML	-0.93	0.84	25.51	0.81	0.88	34.84
	M	-2.51	0.58	12.67	2.96	0.56	15.31

TABLE 4: Model 4 ( $\epsilon = 0.10$ ) and three situations.

		$\hat{E}(\hat{\alpha})$	$B(\hat{\alpha})$	$mse(\hat{\alpha})$	$\hat{E}(\hat{\gamma})$	$B(\hat{\gamma})$	$mse(\hat{\gamma})$
$\alpha = -1$	ML	-0.61	0.39	0.16	0.25	0.38	0.08
	M	-0.71	0.29	0.13	0.33	0.18	0.04
$\alpha = -3$	ML	-0.69	0.77	5.34	0.52	0.82	5.60
	M	-1.39	0.54	2.80	1.56	0.46	2.24
$\alpha = -6$	ML	-0.71	0.88	27.88	0.58	0.91	37.08
	M	-1.70	0.72	18.85	2.15	0.68	21.20

TABLE 1: Pure model (Model 1 where  $\epsilon = 0$ ) and three situations.

		$\hat{E}(\hat{\alpha})$	$B(\hat{\alpha})$	$mse(\hat{\alpha})$	$\hat{E}(\hat{\gamma})$	$B(\hat{\gamma})$	$mse(\hat{\gamma})$
$\alpha = -1$	ML	-1.16	0.16	0.33	0.53	0.31	0.16
	M	-1.19	0.19	0.20	0.54	0.33	0.11
$\alpha = -3$	ML	-5.37	0.79	148.98	6.01	1.08	279.11
	M	-3.88	0.29	8.35	4.00	0.39	14.25
$\alpha = -6$	ML	-6.36	0.06	124.96	7.23	0.08	229.72
	M	-6.25	0.04	69.13	7.08	0.06	122.36

tion becomes less reliable. This is probably due to the shape of the likelihood function, that becomes flat and, therefore, is hard to find its maximum location. The bigger the proportion of contamination the worse the behavior of both estimators, but M-estimators remain consistently closer to the true value than ML-estimators, as expected.

As an application, the image shown in Figure 7 is analyzed. Windows of size  $9 \times 9$  are used to estimate the background parameters over a trajectory that spans from the upper left corner to the lower right corner. This trajectory passes through the three corner reflectors of the image, which are clearly seen as white spots.

Figure 8 shows the result of computing both ML (solid line) and M-estimators (dashes) over  $9 \times 9$  windows in every position of the trajectory in Figure 7. The dotted line shows the background value of  $\alpha$ , and the corner reflectors are located in the positions marked with dash-dots. It is clear that the M-estimator is consistently closer than the ML-estimator to the true value at positions where there is contamination. Besides that, it is always possible to compute M-estimators, whereas ML-estimator fails to converge in 5 out of 45 posi-

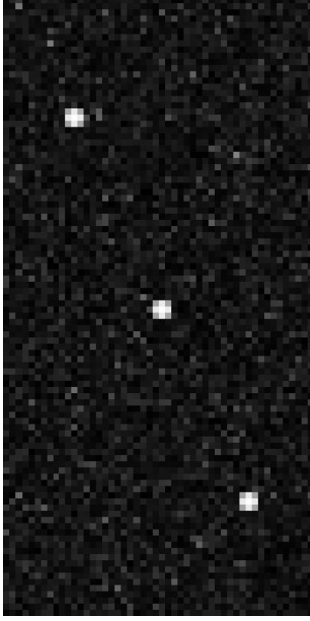


FIGURE 7: Image with three corner reflectors over a heterogeneous background.

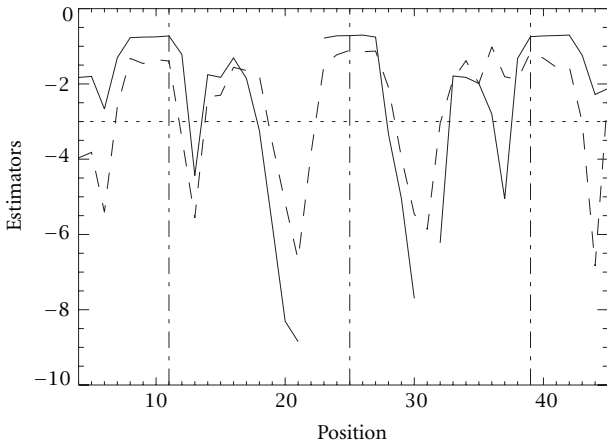


FIGURE 8: Estimated  $\alpha$  using  $9 \times 9$  windows: ML (solid line) and M-(dashes) estimators.

tions (the gaps in the solid line). Similar results were observed estimating the scale parameter  $\gamma$ .

## 6. CONCLUSIONS AND FUTURE WORK

There are numerical problems with the computation of these estimators for certain parameters. For small samples there is often no solution, being this situation more critical for small values of  $\alpha$ , that is, for homogeneous areas. This issue will be further investigated in future works.

As it can be seen from the tables, in the pure model (Model 1), the behavior of the robust estimators is as good as the maximum likelihood ones, as it is expected. In the

contaminated situations (Models 2 and 3), we can see that estimation by moments or maximum likelihood methods is poor.

It is relevant to notice that under a very small contamination or a very small deviation from the model the behavior of the classical estimators is not reliable.

This work will continue computing the M-estimators for the multilook case, that is, for the  $n > 1$  situation and for polarimetric (multivariate) data. An alternative approach using alpha-stable distributions [14] is also possible, but at the expense of losing the interpretability of the parameters that stem from the multiplicative model.

## APPENDICES

### A. COMPUTATION OF $c_1$

Let  $\alpha < 0$ . The problem consists of finding  $c_1(\alpha, 1)$  that satisfies  $I_1(\alpha, c_1, b) = 0$ , where  $I_1(\alpha, c, b) = \int_0^\infty \psi_b((1/\alpha) + \ln(1+u) - c)((-\alpha)/(1+u)^{1-\alpha}) du$  is the equation that has to be solved in order to define the M-estimators for the  $\mathcal{G}_A^0$  distribution (equations (15)).

Lemma A.1. *Let  $\alpha < 0$ , then*

(1) *If  $\exp(2\alpha b) > b\alpha + 1$  then there exists  $c_0$ , with  $-b + \alpha^{-1} < c_0 < b + \alpha^{-1}$  that satisfies  $I_1(\alpha, c_0, b) = 0$ . In this case  $c_0$  is  $c_0 = \exp(\alpha b + \alpha c_0 - 1)$ .*

(2) *If  $\exp(2\alpha b) < b\alpha + 1$  then there exists  $c_0$ , with  $c_0 \geq b + \alpha^{-1}$  that satisfies  $I_1(\alpha, c_0, b) = 0$ , and it is given by  $c_0 = \alpha^{-1}(\ln(-\alpha) + \ln(b) - \ln(e^{-\alpha b} - e^{\alpha b}) + 1)$ .*

*Proof.* The function  $G(c) = I_1(\alpha, c, b)$  is a continuous monotone decreasing function, then

$$\lim_{c \rightarrow -\infty} G(c) = b, \quad \lim_{c \rightarrow \infty} G(c) = -b. \quad (\text{A.1})$$

Therefore, there exists  $c_0$  that satisfies  $G(c_0) = I_1(\alpha, c_0, b) = 0$ . Denoting  $A = A(\alpha, c) = \exp(-b - (1/\alpha) + c) - 1$  and  $B = B(\alpha, c) = \exp(b - \alpha^{-1} + c) - 1$ , then  $A < B$  and one can write

$$\psi_b \left( \frac{1}{\alpha} + \ln(1+u) - c \right) = \begin{cases} -b & \text{if } u < A, \\ \frac{1}{\alpha} + \ln(1+u) - c & \text{if } A < u < B, \\ b & \text{if } u > B. \end{cases} \quad (\text{A.2})$$

If  $c < -b + \alpha^{-1}$ , then  $B < 0$ , and then  $I_1(\alpha, c, b) = b$ . So, there is no  $c_0 < -b + \alpha^{-1}$  satisfying  $I_1(\alpha, c_0, b) = 0$ . If  $-b + \alpha^{-1} \leq c < b + \alpha^{-1}$ , then  $A < 0 < B$  and, therefore,

$$I_1(\alpha, c, b) = \frac{\exp(\alpha b + \alpha c - 1)}{\alpha} - c. \quad (\text{A.3})$$

As  $I_1(\alpha, -b + \alpha^{-1}, b) = b > 0$ , then  $c_0 > -b + \alpha^{-1}$ , and

$$I_1 \left( \alpha, -b + \frac{1}{\alpha}, b \right) = \frac{1}{\alpha} (\exp(2\alpha b) - 1) - b. \quad (\text{A.4})$$

If  $I_1(\alpha, -b + (1/\alpha), b) < 0$ , then there exists  $-b + \alpha^{-1} < c_0 < b + \alpha^{-1}$  such that  $I_1(\alpha, c_0, b) = 0$ . This occurs when  $\exp(2\alpha b) > b\alpha + 1$  and  $c_0$  must satisfy

$$c_0 = \alpha^{-1} \exp(\alpha b_1 + \alpha c_0 - 1). \quad (\text{A.5})$$

If this is not the case, that is if  $\exp(2\alpha b) < b\alpha + 1$ , then  $c$  must be greater than  $b_1 + \alpha^{-1}$ . Thus  $0 < A < B$  and then,

$$\begin{aligned} I_1(\alpha, c, b_1) \\ = \alpha^{-1} (\exp(\alpha c + \alpha b_1 - 1) - \exp(\alpha c - \alpha b_1 - 1)) - b_1. \end{aligned} \quad (\text{A.6})$$

In this situation,  $c_0$  satisfying  $I_1(\alpha, c_0, b) = 0$  is given by  $c_0 = \alpha^{-1} (\ln(-\alpha) + \ln(b) - \ln(e^{-\alpha b} - e^{\alpha b}) + 1)$ .  $\square$

## B. COMPUTATION OF $c_2$

Let  $\alpha < 0$  and  $\gamma > 0$ . As for the computation of  $c_1$ , the problem of finding the function  $c_2$  consists of finding  $c_2 = c_2(\alpha, \gamma, b_2)$  that satisfies

$$I_2(\alpha, \gamma, c_2, b_2) = 0, \quad (\text{B.1})$$

where  $I_2(\alpha, \gamma, c, b_2) = \int_0^\infty \psi_{\gamma b_2}(-\alpha - (1-\alpha)/(1+u) - c)(-\alpha)/(1+u)^{1-\alpha} du$ . In this situation there are more cases than in the computation of  $c_1$  due to the presence of the parameter  $\gamma$ .

**Lemma B.1.** *Let  $\alpha < 0$  and  $\gamma > 0$ . In order to find the function  $c_2$  that satisfies (B.1) one has to consider the following cases:*

- (1) if  $1 - \alpha - 2\gamma b_2 > 0$  then
  - (a) if  $-(1 - \alpha - 2\gamma b_2)^{1-\alpha} \leq (\gamma b_2 - 1)(1 - \alpha)^{1-\alpha}$  then there exists  $c_2$  such that  $c_2 \leq (\gamma b_2 - 1)$ , and it is given by the solution of  $c_2 = -((- \gamma b_2 - \alpha - c_2)/(1 - \alpha))^{1-\alpha}$ ,
  - (b) if  $-(1 - \alpha - 2\gamma b_2)^{1-\alpha} > (\gamma b_2 - 1)(1 - \alpha)^{1-\alpha}$  then  $c_2$  is given by
    - (i) the solution of  $\gamma b_2(1 - \alpha)^{1-\alpha} = (\gamma b_2 - \alpha - c_2)^{1-\alpha} - (-\gamma b_2 - \alpha - c_2)^{1-\alpha}$  if  $(2\gamma b_2/(1 - \alpha))^{1-\alpha} - \gamma b_2 < 0$ , or
    - (ii)  $c_2 = \gamma b_2 - \alpha - (\gamma b_2)^{1/(1-\alpha)}(1 - \alpha)$  if  $(2\gamma b_2/(1 - \alpha))^{1-\alpha} - \gamma b_2 \geq 0$ ,
- (2) if  $1 - \alpha - 2\gamma b_2 \leq 0$  then  $c_2$  satisfying (B.1) is given by
  - (a) the solution of  $c_2 = -((- \gamma b_2 - \alpha - c_2)/(1 - \alpha))^{1-\alpha}$ , if  $\gamma b_2 < -\alpha$ , or
  - (b)  $c_2 = 0$ , if  $\gamma b_2 > -\alpha$  and  $\gamma b_2 > 1$ , or
  - (c)  $c_2 = \gamma b_2 - \alpha - (\gamma b_2)^{(1-\alpha)^{-1}}(1 - \alpha)$  if  $\gamma b_2 > -\alpha$  and  $\gamma b_2 \leq 1$ .

*Proof.* As the function  $G(c) = I_2(\alpha, \gamma, c, b)$  is a continuous monotone decreasing function, and

$$\lim_{c \rightarrow -\infty} G(c) = b_2, \quad \lim_{c \rightarrow \infty} G(c) = -b_2, \quad (\text{B.2})$$

then there exists  $c_2$  that satisfies  $G(c_2) = I_2(\alpha, \gamma, c_2, b_2) = 0$ .

$$\begin{aligned} \psi_{\gamma b_2} \left( -\alpha - \frac{1-\alpha}{1+u} - c \right) \\ = \begin{cases} -\gamma b_2 & \text{if } -\alpha - \frac{1-\alpha}{1+u} - c < -\gamma b_2, \\ -\alpha - \frac{1-\alpha}{1+u} - c & \text{if } -\gamma b_2 < -\alpha - \frac{1-\alpha}{1+u} - c < \gamma b_2, \\ \gamma b_2 & \text{if } -\alpha - \frac{1-\alpha}{1+u} - c > \gamma b_2. \end{cases} \end{aligned} \quad (\text{B.3})$$

*Case 1.* Suppose that  $1 - \alpha - 2\gamma b_2 > 0$ , so  $\gamma b_2 - 1 < -\alpha - \gamma b_2$ . If  $c \leq \gamma b_2 - 1$  and  $A = (1 - \gamma b_2 + c)/(\gamma b_2 - \alpha - c)$ , then  $c < -\alpha - \gamma b_2$ , and

$$\begin{aligned} \psi_{\gamma b_2} \left( -\alpha - \frac{1-\alpha}{1+u} - c \right) \\ = \begin{cases} -\gamma b_2 & \text{if } u < A, \\ -\alpha - \frac{1-\alpha}{1+u} - c & \text{if } A \leq u \leq B, \\ \gamma b_2 & \text{if } u > B, \end{cases} \end{aligned} \quad (\text{B.4})$$

where  $B = (1 + \gamma b_2 + c)/(-\gamma b_2 - \alpha - c)$ ; then

$$I_2(\alpha, \gamma, c, b_2) = - \left( \frac{-\gamma b_2 - \alpha - c}{1 - \alpha} \right)^{1-\alpha} - c. \quad (\text{B.5})$$

Taking  $c = \gamma b_2 - 1$  one has that  $I_2(\alpha, \gamma, \gamma b_2 - 1, b_2) = -(1 - 2\gamma b_2/(1 - \alpha))^{1-\alpha} - (\gamma b_2 - 1)$ , therefore, if  $-(1 - \alpha - 2\gamma b_2)^{1-\alpha} \leq (\gamma b_2 - 1)(1 - \alpha)^{1-\alpha}$  one can say that there exists  $c_2(\alpha, \gamma, b_2)$  satisfying (B.1), given by the solution of  $c_2 = -((- \gamma b_2 - \alpha - c_2)/(1 - \alpha))^{1-\alpha}$ .

For the other situation, if  $-(1 - \alpha - 2\gamma b_2)^{1-\alpha} > (\gamma b_2 - 1)(1 - \alpha)^{1-\alpha}$  one has to look for  $c_2$  in the interval  $(\gamma b_2 - 1, -\alpha - \gamma b_2]$ . In this case,

$$\begin{aligned} I_2(\alpha, \gamma, c, b_2) = -\gamma b_2 + \left( \frac{\gamma b_2 - \alpha - c_2}{1 - \alpha} \right)^{1-\alpha} \\ - \left( \frac{-\gamma b_2 - \alpha - c_2}{1 - \alpha} \right)^{1-\alpha}. \end{aligned} \quad (\text{B.6})$$

Replacing  $c$  in the above equation by  $-\alpha - \gamma b_2$  one obtains  $I_2(\alpha, \gamma, -\alpha - \gamma b_2, b_2) = (2\gamma b_2/(1 - \alpha))^{1-\alpha} - \gamma b_2$ . Then, if  $(2\gamma b_2/(1 - \alpha))^{1-\alpha} - \gamma b_2 \leq 0$ , so there exists  $\gamma b_2 - 1 < c_2 \leq -\alpha - \gamma b_2$  satisfying (B.1) and it is given by the solution  $c_2$  that makes equation (B.6) zero.

But, when  $(2\gamma b_2/(1 - \alpha))^{1-\alpha} - \gamma b_2 > 0$  one has to seek for the solution of (B.1) in the interval  $(-\alpha - \gamma b_2, +\infty)$ . In this case

$$I_2(\alpha, \gamma, c, b_2) = \left( \frac{\gamma b_2 - \alpha - c}{1 - \alpha} \right)^{1-\alpha} - \gamma b_2. \quad (\text{B.7})$$

Hence,  $c_2 = \gamma b_2 - \alpha - (\gamma b_2)^{(1-\alpha)^{-1}}(1 - \alpha)$ .

*Case 2.* Consider the situation  $1 - \alpha - 2\gamma b_2 \leq 0$ , so that  $\gamma b_2 - 1 \geq -\alpha - \gamma b_2$ . If  $c \leq -\alpha - \gamma b_2$  then one has that



$I_2(\alpha, \gamma, c, b_2) = -((- \gamma b_2 - \alpha - c)/(1 - \alpha))^{1-\alpha} - c$ . Taking  $c = -\alpha - \gamma b_2$ , then  $I_2(\alpha, \gamma, -\alpha - \gamma b_2, b_2) = \alpha + \gamma b_2$ . As  $I_2(\alpha, \gamma, -\alpha - \gamma b_2, b_2) < 0$  if  $\gamma b_2 < -\alpha$ , one can say that  $c_2$  satisfying equation (B.1) is given by the solution of  $c_2 = -((- \gamma b_2 - \alpha - c_2)/(1 - \alpha))^{1-\alpha}$ .

But, if  $\gamma b_2 \geq -\alpha$ , one can take  $-\gamma b_2 - \alpha < c \leq \gamma b_2 - 1$ . In this case one has that  $I_2(\alpha, \gamma, c, b_2) = -c$ , therefore  $I_2(\alpha, \gamma, \gamma b_2 - 1, b_2) = 1 - \gamma b_2$ . Thus, if  $1 < \gamma b_2$  one can say that the solution of equation (B.1) is  $c_2 = 0$ .

If  $\gamma b_2 \leq 1$ , then one takes  $c > \gamma b_2 - 1$ , leading to

$$I_2(\alpha, \gamma, c, b_2) = \left( \frac{\gamma b_2 - \alpha - c}{1 - \alpha} \right)^{1-\alpha} - \gamma b_2. \quad (\text{B.8})$$

Hence,  $c_2 = \gamma b_2 - \alpha - (\gamma b_2)^{1/(1-\alpha)}(1 - \alpha)$  is the solution of (B.1).  $\square$

## C. COMPUTATION OF TUNING PARAMETERS

The definition of efficient M-estimators requires finding tuning parameters  $b_1$  and  $b_2$  ensuring that the relative asymptotic variance (with respect to the variance of maximum likelihood estimators) satisfies  $V(\hat{\alpha}_{ML})/V(\hat{\alpha}_M) \cong e_1$  and  $V(\hat{\gamma}_{ML})/V(\hat{\gamma}_M) \cong e_2$  with, for instance,  $0.9 \leq e_i \leq 0.975$ .  $V(\hat{\alpha}_{ML})$ ,  $V(\hat{\gamma}_{ML})$  and  $V(\hat{\alpha}_M)$ ,  $V(\hat{\gamma}_M)$  are the diagonal elements of the asymptotic covariance matrix of the ML- and M-estimators of parameters  $\alpha$  and  $\gamma$ , here denoted by  $V((\hat{\alpha}, \hat{\gamma})_{ML})$  and  $V((\hat{\alpha}, \hat{\gamma})_M)$ , respectively.

### C.1. Computation of $V((\alpha, \gamma)_{ML})$

We denote

$$V((\alpha, \gamma)_{ML}) = M(s, F(\alpha, \gamma))^{-1} J(\alpha, \gamma) M(s, F(\alpha, \gamma))^{-T}, \quad (\text{C.1})$$

where  $F(\alpha, \gamma)$  is the cumulative distribution function of a random variable having  $G_A^0(\alpha, \gamma, 1)$  distribution,  $s = s(x, (\alpha, \gamma))$  is the vector of score functions,

$$J(\alpha, \gamma) = \int s(x, (\alpha, \gamma)) s(x, (\alpha, \gamma))^T dF(x, (\alpha, \gamma)) \quad (\text{C.2})$$

is the Fisher information matrix, and

$$M(s, F(\alpha, \gamma)) = - \int \left[ \frac{\partial}{\partial(\alpha, \gamma)} s(x, (\alpha, \gamma)) \right] dF(x, (\alpha, \gamma)). \quad (\text{C.3})$$

It can be seen that

$$J(\alpha, \gamma) = \begin{bmatrix} \alpha^{-2} & (\gamma(1 - \alpha))^{-1} \\ (\gamma(1 - \alpha))^{-1} & -\alpha((2 - \alpha)\gamma^2)^{-1} \end{bmatrix}, \quad (\text{C.4})$$

and that

$$\begin{aligned} M(s, F(\alpha, \gamma)) &= J(\alpha, \gamma) \\ &= \begin{bmatrix} \alpha^{-2} & (\gamma(1 - \alpha))^{-1} \\ (\gamma(1 - \alpha))^{-1} & -\alpha((2 - \alpha)\gamma^2)^{-1} \end{bmatrix}. \end{aligned} \quad (\text{C.5})$$

Therefore, replacing matrix (C.4) and the inverse of matrix (C.5) in equation (C.1), we have

$$\begin{aligned} V((\alpha, \gamma)_{ML}) &= \begin{bmatrix} \alpha^2(1 - \alpha)^2 & \gamma\alpha(1 - \alpha)(2 - \alpha) \\ \gamma\alpha(1 - \alpha)(2 - \alpha) & -\alpha^{-1}\gamma^2(1 - \alpha)^2(2 - \alpha) \end{bmatrix}. \end{aligned} \quad (\text{C.6})$$

### C.2. Computation of $V((\alpha, \gamma)_M)$

We define

$$V((\alpha, \gamma)_M) = M(\Psi, F(\alpha, \gamma))^{-1} Q(\alpha, \gamma) M(\Psi, F(\alpha, \gamma))^{-T}, \quad (\text{C.7})$$

where

$$\begin{aligned} \Psi &= \Psi(x; (\alpha, \gamma)) = \begin{bmatrix} \psi_{b_1}(s_1(x; (\alpha, \gamma)) - c_1(\alpha, \gamma, b_1)) \\ \psi_{b_2}(s_2(x; (\alpha, \gamma)) - c_2(\alpha, \gamma, b_2)) \end{bmatrix}, \\ M(\Psi, F(\alpha, \gamma)) &= - \frac{\partial}{\partial \alpha} \left[ \int \Psi(x, (\alpha, \gamma)) dF(x, (\alpha, \gamma)) \right], \\ Q(\alpha, \gamma) &= \int \Psi(x, (\alpha, \gamma)) \Psi(x, (\alpha, \gamma))^T dF(x, (\alpha, \gamma)). \end{aligned} \quad (\text{C.8})$$

The matrices  $M$  and  $Q$  are computed taking into account all the cases that define the functions  $c_i$ ,  $i = 1, 2$ , and solving explicitly the integrals.

## ACKNOWLEDGEMENT

The authors are grateful to Conicor, SeCyT, CNPq, and Vitae for the partial support of this research.

## REFERENCES

- [1] C. Oliver and S. Quegan, *Understanding Synthetic Aperture Radar Images*, Artech House, Boston, 1998.
- [2] A. C. Frery, H.-J. Müller, C. C. F. Yanasse, and S. J. S. Sant'Anna, "A model for extremely heterogeneous clutter," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35, pp. 648–659, 1997.
- [3] J. Jacobo-Berlles, M. Mejail, and A. C. Frery, "The GA0 distribution as the true model for SAR images," in *Simpósio Brasileiro de Computação Gráfica e Processamento de Imagens*, SBC, IEEE, Campinas, SP, Brazil, 1999.
- [4] M. Mejail, J. Jacobo-Berlles, A. C. Frery, and O. H. Bustos, "Parametric roughness estimation in amplitude SAR images under the multiplicative model," *Revista de Teledetección*, vol. 13, pp. 37–49, 2000.
- [5] M. E. Mejail, A. C. Frery, J. Jacobo-Berlles, and O. H. Bustos, "Approximation of distributions for SAR images: Proposal, evaluation and practical consequences," *Latin American Applied Research*, vol. 31, pp. 83–92, 2001.
- [6] O. H. Bustos, "Robust statistics in SAR image processing," in *Proceedings of the First Latino-American Seminar on Radar Remote Sensing: Image Processing Techniques*, pp. 81–89, ESA, Paris, 1996.
- [7] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel, *Robust Statistics: The Approach Based on Influence Functions*, Wiley, New York, 1986.
- [8] A. Marazzi and C. Ruffieux, "Implementing M-estimators of the gamma distribution," in *Robust Statistics, Data Analysis*,

- and *Computer Intensive Methods*, R. Helmut, Ed., vol. 109 of *Lecture Notes in Statistics*, Springer-Verlag, Berlin, 1996.
- [9] A. C. Frery, A. H. Correia, C. D. Rennó, C. C. Freitas, J. Jacobo-Berlles, M. E. Mejail, and K. L. P. Vasconcellos, "Models for synthetic aperture radar image analysis," *Resenhas (IME-USP)*, vol. 4, no. 1, pp. 45–77, 1999.
- [10] O. E. Barndorff-Nielsen and P. Blæsild, "Hyperbolic distributions and ramifications: Contributions to theory and applications," in *Statistical distributions in scientific work*, C. Taillie and B. A. Baldessari, Eds., pp. 19–44, Reidel, Dordrecht, 1981.
- [11] A. C. Frery, C. C. F. Yanasse, and S. J. S. Sant'Anna, "Alternative distributions for the multiplicative model in SAR images," in *Quantitative remote sensing for science and applications*, IGARSS'95 Proc., pp. 169–171, IEEE, Florence, July 1995.
- [12] P. K. Sen and J. M. Singer, *Large Sample Methods in Statistics: An Introduction with Applications*, Chapman and Hall, London, 1993.
- [13] C. F. Manski, *Analog Estimation Methods in Econometrics*, vol. 39 of *Monographs on Statistics and Applied Probability*, Chapman & Hall, New York, 1988, available at <http://elsa.berkeley.edu/books/analog.html>.
- [14] C. L. Nikias and M. Shao, *Signal Processing with Alpha-Stable Distributions and Applications*, Wiley, New York, 1995.

**Oscar H. Bustos** received the B.S. degree in Mathematics from the Universidad Nacional de Córdoba, Argentina, in 1970 and the Ph.D. in Mathematics from the Universidad Nacional de San Luis, Argentina, in 1978. He is Professor at the Universidad Nacional de Córdoba (Mathematics, Physics and Astronomy Institute) and his research areas include mathematical and statistical models for image processing, robustness in statistics and stochastic simulation.



**María Magdalena Lucini** received the B.S. degree in Mathematics from the Universidad Nacional de Córdoba (Mathematics, Physics and Astronomy Institute) in 1996. She is currently pursuing the Ph.D program in Mathematics at the same university. Her research interests are statistical signal processing with emphasis in robust statistics, estimation and filtering theory and applications in signal and image processing for imaging under severe noise.



**Alejandro C. Frery** received the degree in Electronic Engineering from the Universidad de Mendoza, Argentina, in 1987. His MSc degree was in Applied Mathematics (Statistics) from the Instituto de Matemática Pura e Aplicada (IMPA, Rio de Janeiro, Brazil, 1990) and his Ph.D. was in Applied Computing from the Instituto Nacional de Pesquisas Espaciais (INPE, São José dos Campos, Brazil, 1994). He is currently a lecturer the Computer Science Institute of the Universidade Federal de Pernambuco, Brazil. His research areas include stochastic models for image formation and analysis, as well as stochastic simulation.

