

# Mimic Capacity Of Fisher And Generalized Gamma Distributions For High Resolution SAR Image Statistical Modeling

Hélène Sportouche, Jean-Marie Nicolas, and Florence Tupin, *Senior Member, IEEE*

**Abstract**—The aim of this paper is to compare the potential of two popular flexible laws, the Fisher distribution and the Generalized Gamma distribution, for the statistical modeling of high-resolution SAR data through an original “mimicking-based” approach. The presented study allows to evaluate the ability of both laws to correctly imitate or “mimic” another reference law, frequently used for modeling the intensity of SAR images and chosen for instance as the  $\mathcal{K}$  law or the Weibull, Beta or log-normal laws in this work. This study uses log-cumulant statistics for parameter estimation of the imitating law and involves quantitative criteria of comparison based on the Kullback-Leibler divergences between the reference law and the Fisher law or the Generalized Gamma law. The mimicking capacities of both distributions are first analyzed for some sets of parameters describing different studied cases, covering a wide set of possible mimicked reference laws. The high modeling potential of both distributions is then illustrated on heterogeneous subscenes from real SAR intensity data. Pragmatical considerations are also taken into account to draw up recommendations about the preferential use of a distribution and to highlight complementarities of both Fisher and Generalized Gamma distributions, along with limitations of the approach.

**Index Terms**—Statistical modeling, mimic, high-resolution (HR) SAR data, Fisher distribution, Generalized Gamma distribution, K distribution, Beta distribution, Weibull distribution, log-normal distribution, Kullback-Leibler divergence, log-cumulants.

## I. INTRODUCTION

THE modeling of intensity distributions of SAR data is essential for several aims in remote sensing, such as image denoising [1], image segmentation [2], image classification [3], object detection [4] or 3D reconstruction [5], and change detection [6] [7]. A wide variety of laws [8] [9] [10] [11] - among which the exponential law, the Gamma law, the Log-Normal law, the  $\mathcal{K}$  law, Weibull law, or the Beta law - have been employed in the literature for this purpose, depending on the application framework and on the kind of processed SAR data. Nevertheless, it remains difficult to choose the good statistical model for intensity SAR image processing.

H. Sportouche\* is with Altersis, 13851 Aix-en-Provence, France as Research and Development Engineer. This work has been done during her post doctoral stay at LTCL, CNRS, Télécom ParisTech, Université Paris-Saclay, 75013, Paris, France (e-mail: helene.sportouche@telecom-paristech.org).

J.-M. Nicolas and F. Tupin are with LTCL, CNRS, Télécom ParisTech, Université Paris-Saclay, 75013, Paris, France (e-mail: florence.tupin@telecom-paristech.fr; jean-marie.nicolas@telecom-paristech.fr).

Manuscript received September 15, 2016. Manuscript revised March, 2017.

The classical laws described by two parameters - such as the traditional Gamma laws - have proved to be insufficient in some specific cases, especially for the new generation of high-resolution (HR) and very high-resolution (VHR) SAR sensors. This has thus motivated the use of more sophisticated laws described by three parameters or more to overcome these limitations.

Indeed, one of the interest of a “generic” distribution is its ability to describe a large set of homogeneous and heterogeneous areas in an image (for instance ground, vegetation, urban areas, sea, etc.). Instead of a dictionary of distributions dedicated to different behaviors, the use of a well chosen generic distribution simplifies the processing chain. The selection of such a distribution remains an open question.

For SAR images, two flexible laws, the Fisher<sup>1</sup> distribution [2] [3] [9] [10] [13] [14] and the Generalized Gamma distribution [15] [16] [17] [18] [19], are good candidates to be generic distributions<sup>2</sup>. These two laws, described by three parameters, are quite popular in the literature. In [2] and [3], the Fisher distribution is used respectively for image partitioning and image classification especially on HR SAR data in urban areas. In [17] and [19], it is the Generalized Gamma distribution which is exploited for histogram fitting on SAR data in urban, agricultural and mountainous areas.

But, how good are comparatively Fisher and Generalized Gamma distributions to fit different kinds of behaviors? We propose in this paper to compare their potential through an original “mimicking-based” approach [22], [23]. We study their ability to correctly “mimic”<sup>3</sup> a third distribution, which we have chosen for instance in this work as a  $\mathcal{K}$  law or a Weibull or Beta law. Indeed, the  $\mathcal{K}$ , Weibull or Beta laws (described by two or three parameters) have been until now frequently used for the modeling of intensity of SAR imagery [24] [25] [11], [10] and they can thus be considered as examples of reference laws. Besides, given that Fisher and Generalized Gamma distributions can be themselves consid-

<sup>1</sup>The Fisher distribution is also called the second kind Beta law [10], the type VI Pearson law [10] or the  $\mathcal{G}^0$  distribution [12].

<sup>2</sup>The so-called  $\mathcal{G}$  distribution is another interesting generic distribution [12] [20]. However, the probability density function of this distribution is complex and this fact limits its analytical manipulation and, consequently, its application. An initial analysis based on it is proposed in [21].

<sup>3</sup>In this paper, the “mimic” word refers to the ability of a distribution to correctly imitate or copy another one in terms of probability density function fitting in a data modeling framework.

ered as reference laws, it is also interesting to investigate how Fisher and Generalized Gamma are able to mimic each other.

The presented study uses log-cumulant statistics for parameter estimation of the imitating law and introduces quantitative criteria of comparison based on the Kullback-Leibler divergence to analyze the capacities of both distributions to mimic the reference law. We are thus led to derive the analytical expressions of Kullback-Leibler divergences (KLD) between the  $\mathcal{K}$  law and the Fisher law or the Generalized Gamma law, as well as those of Kullback-Leibler divergences (KLD) between the Fisher law and the Generalized Gamma law and reciprocally. It represents a theoretical contribution of this paper (such expressions can be indeed used for multiple purposes in statistical image processing). We are also led to evaluate imitation performances of both distributions on some sets of parameters (describing different cases of mimicked  $\mathcal{K}$  reference laws). To extend this study we also analyze the case of Weibull, Beta and log-normal laws. We finally illustrate the high modeling potential of both Fisher and Generalized Gamma distributions on real SAR subscenes. We draw up recommendations about the preferential use of a distribution in particular which represents an other contribution of this paper.

We will use the method of log-cumulants (MLC)<sup>4</sup> to estimate Fisher and Generalized Gamma parameters. This choice is justified by its numerous practical advantages compared to the classical maximum likelihood method (MLM) or the method of moments (MM) [2] [28] [29]. For the estimation of Generalized Gamma parameters, MLC is easier, faster and more stable than MLM and MM, and offers very competitive performances, in particular for small sample sizes [28] [29]. For the estimation of Fisher parameters, MLC is easier and faster than MLM and provides better global performances than MM, in addition to some other practical benefits [2] ((i) no applicability restriction relative to moment existence has to be considered with MLC, unlike with MM; (ii) for a given set of parameters, the percentage of generated sample realizations for which the estimated parameters are valid, i.e. positive, called "acceptance rate" in [2], is higher with MLC than with MM, especially for small sample sizes).

Besides, let us also underline that the works presented in this paper provides an enhanced content with respect to the ones in [22] and [23], where the mimicking concept has been previously introduced. More precisely, we propose here: (i) to detail the log-cumulant based mimicking methodology briefly presented in [22] and [23] and to demonstrate the simplified analytical expressions of the Kullback-Leibler divergences involved in the comparison criteria; (ii) to illustrate on several synthetic datasets the mimicking performances of Fisher and Generalized Gamma distributions and to carry out an in-depth analysis of the results. We consider here not only  $\mathcal{K}$  laws as reference ones but also alternatively Weibull, Beta and log-normal laws for instance. This allows us to generalize to a wider set of mimicked laws the proposed approach. (iii) to illustrate the intensity modeling potential of both

distributions of interest on real SAR data and to provide precise and additional pragmatical considerations, that can be useful when dealing with modeling law choice. More precisely, we discuss applicability conditions for MLC (relative to log-cumulant locations), numerical instability risks and numerical precision limitations, in order to raise awareness to practical difficulties that can be encountered even when using flexible laws.

The paper is organized as follows: In section II, the analytical probability density functions (pdfs) and the log-cumulant expressions of Fisher, Generalized Gamma and  $\mathcal{K}$  distributions are reminded. In section III, the log-cumulant and Kullback-Leibler divergence based methodology proposed for mimicking ability comparison is presented for the reference  $\mathcal{K}$  law. In section IV, the results obtained on six studied cases corresponding to  $\mathcal{K}$  mimicked examples are in-depth discussed according to two well-defined points (named as "precision" and "robustness" in this paper). In section V, mimicking performances of Fisher and Generalized Gamma distributions to imitate Weibull, Beta and log-normal references law are more briefly analyzed. In section VI, the approach is applied to the mimicking of Fisher by Generalized Gamma and vice versa. In section VII, the interest of using Fisher or Generalized Gamma law for the intensity modeling of real SAR data is illustrated on two examples with different heterogeneity levels and recommendations, jointly based on theoretical and practical aspects, are given to suggest the preferential use of a particular distribution and also to highlight some critical cases. In section VIII, a summary of the main conclusions is drawn and some precise outlooks are envisaged.

## II. PRELIMINARIES

### A. Presentation of the Fisher distribution

The Fisher probability density function (pdf)  $\mathcal{F}_{pdf}(\mu_{\mathcal{F}}, L_{\mathcal{F}}, M_{\mathcal{F}})$  is defined by [9] [10]:

$$\mathcal{F}_{pdf}(x) = \frac{L_{\mathcal{F}}}{M_{\mathcal{F}}\mu_{\mathcal{F}}} \frac{\Gamma(L_{\mathcal{F}} + M_{\mathcal{F}})}{\Gamma(L_{\mathcal{F}})\Gamma(M_{\mathcal{F}})} \frac{\left(\frac{L_{\mathcal{F}}x}{M_{\mathcal{F}}\mu_{\mathcal{F}}}\right)^{L_{\mathcal{F}}-1}}{\left(1 + \frac{L_{\mathcal{F}}x}{M_{\mathcal{F}}\mu_{\mathcal{F}}}\right)^{L_{\mathcal{F}}+M_{\mathcal{F}}}} \quad (1)$$

where  $\Gamma(x)$  is the Gamma function,  $\mu_{\mathcal{F}}$  is a mean parameter and  $L_{\mathcal{F}}$  and  $M_{\mathcal{F}}$  are two shape parameters, weighting respectively the heavy head behavior and the heavy tail behavior of the Fisher distribution. Indeed,  $\mathcal{F}_{pdf}(\mu_{\mathcal{F}}, L_{\mathcal{F}}, M_{\mathcal{F}})$  can also be defined as the Mellin convolution [30] [10] between a Gamma distribution  $\mathcal{G}_{pdf}(\mu_{\mathcal{F}}, L_{\mathcal{F}})$  and an Inverse of Gamma distribution  $\mathcal{IG}_{pdf}(1, M_{\mathcal{F}})$ <sup>5</sup>, which allows us to give an intuitive interpretation to each parameter.

The three first log-cumulants of the Fisher pdf are respectively defined by [9] [10]:

$$\begin{aligned} \widetilde{\kappa}_{1\mathcal{F}_{pdf}} &= \ln(\mu_{\mathcal{F}}) + \Psi(L_{\mathcal{F}}) - \ln(L_{\mathcal{F}}) - \Psi(M_{\mathcal{F}}) + \ln(M_{\mathcal{F}}) \\ \widetilde{\kappa}_{2\mathcal{F}_{pdf}} &= \Psi(1, L_{\mathcal{F}}) + \Psi(1, M_{\mathcal{F}}) \\ \widetilde{\kappa}_{3\mathcal{F}_{pdf}} &= \Psi(2, L_{\mathcal{F}}) - \Psi(2, M_{\mathcal{F}}) \end{aligned} \quad (2)$$

<sup>4</sup>The method of log-cumulants has been introduced in [9] and [10]. In [26] and [27], studies combining the method of log-cumulants and the squared Mahalanobis distance are proposed in order to obtain a parameter estimation method and a goodness of fit method based on the Mellin transform.

<sup>5</sup>The Fisher distribution and the  $\mathcal{G}^0$  distribution were proposed with the same basic idea, considering both of them a Gamma distribution and an Inverse of Gamma distribution in a multiplicative model context. The difference relies only on a re-parametrization process [10] [12].

where  $\Psi(x)$  and  $\Psi(n, x)$  represent the Digamma function and the  $n$ th Polygamma function respectively.

### B. Presentation of the Generalized Gamma distribution

The Generalized Gamma probability density function (pdf)  $\mathcal{G}\mathcal{G}_{pdf}(\mu_{\mathcal{G}\mathcal{G}}, L_{\mathcal{G}\mathcal{G}}, \eta_{\mathcal{G}\mathcal{G}})$  is defined by [10] [19]:

$$\mathcal{G}\mathcal{G}_{pdf}(x) = \frac{|\eta_{\mathcal{G}\mathcal{G}}| L_{\mathcal{G}\mathcal{G}}^{\left(\frac{1}{\eta_{\mathcal{G}\mathcal{G}}}\right)} \Gamma\left(\frac{1}{\eta_{\mathcal{G}\mathcal{G}}}\right)}{\mu_{\mathcal{G}\mathcal{G}} \Gamma(L_{\mathcal{G}\mathcal{G}})} \left(\frac{L_{\mathcal{G}\mathcal{G}}}{\mu_{\mathcal{G}\mathcal{G}}}\right)^{\eta_{\mathcal{G}\mathcal{G}} L_{\mathcal{G}\mathcal{G}} - 1} \times \exp\left(-\left(\frac{L_{\mathcal{G}\mathcal{G}}}{\mu_{\mathcal{G}\mathcal{G}}}\right)^{\eta_{\mathcal{G}\mathcal{G}}} x\right) \quad (3)$$

where  $\Gamma(x)$  is the Gamma function,  $\mu_{\mathcal{G}\mathcal{G}}$  is a mean parameter,  $L_{\mathcal{G}\mathcal{G}}$  is a shape parameter and  $\eta_{\mathcal{G}\mathcal{G}}$  is a power parameter.

The three first log-cumulants of the Generalized Gamma pdf are respectively defined by [10] [19]:

$$\begin{aligned} \widetilde{\kappa}_1 \mathcal{G}\mathcal{G}_{pdf} &= \ln(\mu_{\mathcal{G}\mathcal{G}}) + \frac{\Psi(L_{\mathcal{G}\mathcal{G}}) - \ln(L_{\mathcal{G}\mathcal{G}})}{\eta_{\mathcal{G}\mathcal{G}}} \\ \widetilde{\kappa}_2 \mathcal{G}\mathcal{G}_{pdf} &= \frac{\Psi(1, L_{\mathcal{G}\mathcal{G}})}{\eta_{\mathcal{G}\mathcal{G}}^2} \\ \widetilde{\kappa}_3 \mathcal{G}\mathcal{G}_{pdf} &= \frac{\Psi(2, L_{\mathcal{G}\mathcal{G}})}{\eta_{\mathcal{G}\mathcal{G}}^3} \end{aligned} \quad (4)$$

### C. Presentation of the $\mathcal{K}$ distribution

The  $\mathcal{K}$  probability density function (pdf)  $\mathcal{K}_{pdf}(\mu_{\mathcal{K}}, L_{\mathcal{K}}, M_{\mathcal{K}})$  is defined by [9] [10]:

$$\mathcal{K}_{pdf}(x) = \frac{2L_{\mathcal{K}}M_{\mathcal{K}}}{\mu_{\mathcal{K}}} \frac{1}{\Gamma(L_{\mathcal{K}})\Gamma(M_{\mathcal{K}})} \left(\frac{L_{\mathcal{K}}M_{\mathcal{K}}x}{\mu_{\mathcal{K}}}\right)^{\frac{L_{\mathcal{K}}+M_{\mathcal{K}}}{2}-1} \times \mathcal{BK}\left(M_{\mathcal{K}} - L_{\mathcal{K}}, 2\left(\frac{L_{\mathcal{K}}M_{\mathcal{K}}x}{\mu_{\mathcal{K}}}\right)^{\frac{1}{2}}\right) \quad (5)$$

where  $\Gamma(x)$  is the Gamma function,  $\mathcal{BK}(n, x)$  is the second kind modified Bessel function of order  $n$ ,  $\mu_{\mathcal{K}}$  is a mean parameter and  $L_{\mathcal{K}}$  and  $M_{\mathcal{K}}$  are two shape parameters. In a similar way to the Fisher pdf,  $\mathcal{K}_{pdf}(\mu_{\mathcal{K}}, L_{\mathcal{K}}, M_{\mathcal{K}})$  can also be defined as the Mellin convolution [30] [10] between a Gamma distribution  $\mathcal{G}_{pdf}(\mu_{\mathcal{K}}, L_{\mathcal{K}})$  and an other Gamma distribution  $\mathcal{G}_{pdf}(1, M_{\mathcal{K}})$ .

The three first log-cumulants of the  $\mathcal{K}$  pdf are respectively defined by [9] [10]:

$$\begin{aligned} \widetilde{\kappa}_1 \mathcal{K}_{pdf} &= \ln(\mu_{\mathcal{K}}) + \Psi(L_{\mathcal{K}}) - \ln(L_{\mathcal{K}}) + \Psi(M_{\mathcal{K}}) - \ln(M_{\mathcal{K}}) \\ \widetilde{\kappa}_2 \mathcal{K}_{pdf} &= \Psi(1, L_{\mathcal{K}}) + \Psi(1, M_{\mathcal{K}}) \\ \widetilde{\kappa}_3 \mathcal{K}_{pdf} &= \Psi(2, L_{\mathcal{K}}) + \Psi(2, M_{\mathcal{K}}) \end{aligned} \quad (6)$$

The special cases of  $\mathcal{K}$  distributions where  $L_{\mathcal{K}} = M_{\mathcal{K}}$  are named “ $\mathcal{K}_c$  caustic distributions” in this paper. The  $\mathcal{K}_c$  probability density function (pdf) is thus defined by:

$$\mathcal{K}_c pdf(x) = \frac{2L_{\mathcal{K}_c}^2}{\mu_{\mathcal{K}_c}} \frac{1}{\Gamma(L_{\mathcal{K}_c})^2} \left(\frac{L_{\mathcal{K}_c}^2 x}{\mu_{\mathcal{K}_c}}\right)^{L_{\mathcal{K}_c} - 1} \times \mathcal{BK}\left(0, 2\left(\frac{L_{\mathcal{K}_c}^2 x}{\mu_{\mathcal{K}_c}}\right)^{\frac{1}{2}}\right) \quad (7)$$

where  $\Gamma(x)$  is the Gamma function,  $\mathcal{BK}(n, x)$  is the second kind modified Bessel function of order  $n$ ,  $\mu_{\mathcal{K}_c}$  is a mean parameter and  $L_{\mathcal{K}_c}$  is a shape parameter.

### D. Other usual SAR distributions

Among other distributions widely used in SAR imagery, we have chosen the Weibull distribution, the Beta distribution and the log-normal distribution to illustrate the mimicking capacities of the Generalized Gamma and Fisher distributions.

The Weibull pdf (with 2 parameters  $\mu_{\mathcal{W}}$  and  $\eta_{\mathcal{W}}$ ) is given by:

$$\mathcal{W}_{pdf}(x) = \frac{|\eta_{\mathcal{W}}|}{\mu_{\mathcal{W}}} \left(\frac{x}{\mu_{\mathcal{W}}}\right)^{\eta_{\mathcal{W}} - 1} e^{-\left(\frac{x}{\mu_{\mathcal{W}}}\right)^{\eta_{\mathcal{W}}}} \quad (8)$$

As it can be seen, it is a special case of a Generalized Gamma pdf with  $L_{\mathcal{G}\mathcal{G}} = 1$  in eq.3.

The Beta pdf (three parameters  $\mu_{\mathcal{B}}, L_{\mathcal{B}}, M_{\mathcal{B}}$ ) is given by:

$$\mathcal{B}_{pdf}(x) = \frac{L_{\mathcal{B}}}{M_{\mathcal{B}}\mu_{\mathcal{B}}} \frac{\Gamma(M_{\mathcal{B}})}{\Gamma(L_{\mathcal{B}})\Gamma(M_{\mathcal{B}}-L_{\mathcal{B}})} \left(\frac{L_{\mathcal{B}}x}{M_{\mathcal{B}}\mu_{\mathcal{B}}}\right)^{L_{\mathcal{B}}-1} \left(1 - \frac{L_{\mathcal{B}}x}{M_{\mathcal{B}}\mu_{\mathcal{B}}}\right)^{M_{\mathcal{B}}-L_{\mathcal{B}}-1} \quad x \in \left[0; \frac{M_{\mathcal{B}}\mu_{\mathcal{B}}}{L_{\mathcal{B}}}\right] \quad (9)$$

The log-normal pdf (two parameters  $\mu_{\mathcal{L}}, \sigma_{\mathcal{L}}$ ) is given by:

$$\mathcal{L}_{pdf}(x) = \frac{1}{\sigma_{\mathcal{L}}\sqrt{2\pi}x} e^{-\left(\frac{(\log x - \mu_{\mathcal{L}})^2}{2\sigma_{\mathcal{L}}^2}\right)} \quad (10)$$

### E. Representation of these distributions in the $\widetilde{\kappa}_2 - \widetilde{\kappa}_3$ diagram

The so-called  $\widetilde{\kappa}_2 - \widetilde{\kappa}_3$  diagram, whose use is well-adapted when dealing with second kind statistics, allows a simple characterization of the different laws [9] [10] [31]. The positions of the previously described distributions are presented in this diagram in figure 1.

The Fisher distributions (described by three parameters) correspond to the area located above the branches of curves designating the Gamma and Inverse of Gamma distributions (described by two parameters).

In order to apply a log-cumulant based method for Fisher parameter estimation, the condition defined by the following equation has thus to be satisfied [28] [29]:

$$\widehat{\kappa}_2 \mathcal{F}_{pdf} \geq \Psi\left(1, \Phi\left(2, -\left|\widehat{\kappa}_3 \mathcal{F}_{pdf}\right|\right)\right) \quad (11)$$

where  $\Psi(n, x)$  and  $\Phi(n, x)$  represent the  $n$ th Polygamma function and the inverse of the  $n$ th Polygamma function respectively and  $\widehat{\kappa}_2 \mathcal{F}_{pdf}$  and  $\widehat{\kappa}_3 \mathcal{F}_{pdf}$  are the empirical log-cumulants of the Fisher distribution.

The Generalized Gamma distributions (described by three parameters) correspond to the larger area located above the branches of curves defined by the following equation [28] [29], excluding the ordinate axis (designating the log-normal distributions):

$$\frac{\widetilde{\kappa}_3 \mathcal{G}\mathcal{G}_{pdf}}{\widetilde{\kappa}_2 \mathcal{G}\mathcal{G}_{pdf}} = \frac{1}{4} \quad (12)$$

In order to apply a log-cumulant based method for Generalized Gamma parameter estimation, the condition defined by the following equation has thus to be satisfied [28] [29]:

$$\frac{\widehat{\kappa}_3 \mathcal{G}\mathcal{G}_{pdf}}{\widehat{\kappa}_2 \mathcal{G}\mathcal{G}_{pdf}} > \frac{1}{4} \quad (13)$$

where  $\widehat{\kappa}_2 \mathcal{G}\mathcal{G}_{pdf}$  and  $\widehat{\kappa}_3 \mathcal{G}\mathcal{G}_{pdf}$  are the empirical log-cumulants of the Generalized Gamma distribution.

The generic characteristic of Fisher and Generalized Gamma can be seen in the  $\tilde{\kappa}_2 - \tilde{\kappa}_3$  diagram by the large areas they are both covering in this space.

The  $\mathcal{K}$  distributions (described by three parameters) correspond to the smaller area located above the branch of the curve designating the Gamma distributions and below the branch of the  $\mathcal{K}_c$  caustic curve designating the special cases of  $\mathcal{K}$  distributions where  $L_{\mathcal{K}} = M_{\mathcal{K}}$ .

Let us denote that Fisher distributions and  $\mathcal{K}$  distributions share a same border curve, i.e. the branch of curve designating the Gamma distribution, because it can be mathematically shown that [10]: given the Mellin convolution-based definitions of Fisher laws and  $\mathcal{K}$  laws, when  $M_{\mathcal{F}}$  tends towards the Infinite,  $\mathcal{F}_{pdf}(\mu_{\mathcal{F}}, L_{\mathcal{F}}, M_{\mathcal{F}})$  tends towards  $\mathcal{G}_{pdf}(\mu_{\mathcal{F}}, L_{\mathcal{F}})$  and when  $M_{\mathcal{K}}$  tends towards the Infinite,  $\mathcal{K}_{pdf}(\mu_{\mathcal{K}}, L_{\mathcal{K}}, M_{\mathcal{K}})$  tends towards  $\mathcal{G}_{pdf}(\mu_{\mathcal{K}}, L_{\mathcal{K}})$ .

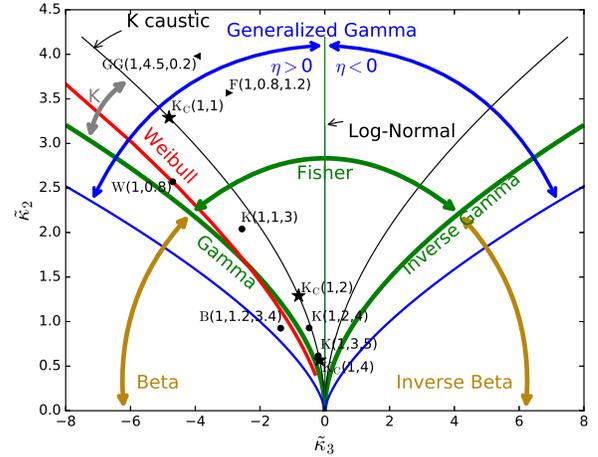
The  $\mathcal{K}$  distributions can thus be theoretically mimicked by a Fisher law as well as by a Generalized Gamma law. The Weibull distribution is a special case of Generalized Gamma distribution and can be mimicked by a Fisher distribution. The log-normal distribution belongs to the domain of Fisher pdf, but corresponds to a discontinuity for Generalized Gamma. Finally, the domain of Beta distributions partly overlaps with Generalized Gamma domain but not with Fisher one.

### III. PROPOSED METHOD TO COMPARE FISHER AND GENERALIZED GAMMA ABILITY TO MIMICK A REFERENCE $\mathcal{K}$ LAW

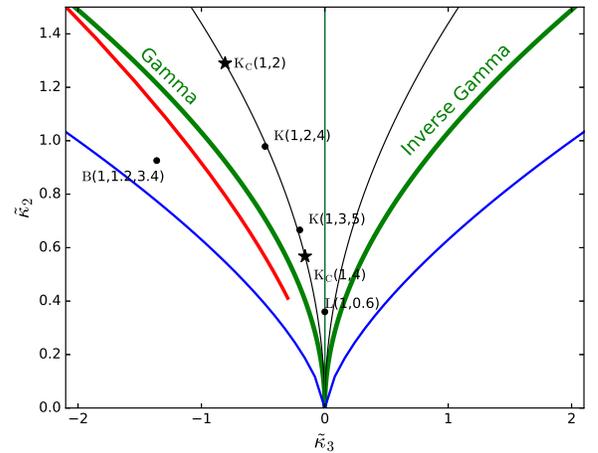
In this section, a two-stage methodology is proposed to compare the ability of Fisher and Generalized Gamma to mimic a reference  $\mathcal{K}$  law. First, the parameters of the mimicking laws are estimated using the MLC. Second, a qualitative visual criterion and some quantitative Kullback-Leibler divergence-based criteria are computed to lead a comparison.

Two cases are distinguished in this study: (i) case 1: we assume that the parameters of the reference mimicked  $\mathcal{K}$  law are known; (ii) case 2: we assume that the parameters of the reference mimicked  $\mathcal{K}$  law are unknown, like this is the case in practice, but that sample values distributed according to this  $\mathcal{K}$  law are available (they correspond, for instance on real SAR data, to some intensity values of pixels belonging to an area, whose histogram is supposed to be quite well modeled by a  $\mathcal{K}$  law).

In all this work we will only focus on the shape parameters controlling the distribution. As it has been seen previously, for all the considered pdf, these parameters are given by the  $\tilde{\kappa}_2$  and  $\tilde{\kappa}_3$  log-cumulants. After the estimation of the shape parameters, the exploitation of  $\tilde{\kappa}_1$  provides an estimation of the mean parameter  $\mu_{pdf}$ . In theory, the  $\mu_{pdf}$  value can have an impact on the mimicking potential of a distribution since it is also estimated with the shape parameters. Nevertheless empirical studies showed that the results for the same set of parameters except for the  $\mu_{pdf}$  value for mimicking / mimicked laws give very similar results to the case where  $\mu_{pdf}$  equal 1. Therefore in all the studies of this paper, the mean parameter will be chosen equal to 1 for the pdf to be mimicked. Besides, part VII show results on real SAR images with good histogram fitting whatever the  $\mu$  values.



(a)  $\tilde{\kappa}_2 - \tilde{\kappa}_3$  diagram



(b) Zoom

Fig. 1. (a) Representation of the Fisher, Generalized Gamma, Beta,  $\mathcal{K}$ , Weibull and log-normal distribution locations in the  $\tilde{\kappa}_2 - \tilde{\kappa}_3$  diagram. The stars and the bullets represent the location of the points associated to values used in section IV for the six studied sets of parameters, respectively on the  $\mathcal{K}_c$  caustic curve and in the  $\mathcal{K}$  area. The triangles represent the location of the points associated to values used in section VI for the two studied sets of parameters in the Fisher area (common to Fisher and Generalized Gamma). The Weibull, Beta and log-normal pdf studied in part V are also indicated (on the top for Weibull pdf, on the top and bottom for Beta pdf and on the bottom for log-normal pdf). (b) Zoom on a sub-part of the  $\tilde{\kappa}_2 - \tilde{\kappa}_3$  diagram.

#### A. Log-cumulant based estimation of mimicking law parameters

In the case 1, the log-cumulant based parameter estimation of the mimicking law (here, the Fisher law or the Generalized Gamma law), likely to imitate the reference  $\mathcal{K}$  law, can be directly done using the three first true log-cumulants of the  $\mathcal{K}$  distribution through the following steps:

- 1) Compute the true log-cumulants  $\tilde{\kappa}_2 \mathcal{K}_{pdf}$ ,  $\tilde{\kappa}_3 \mathcal{K}_{pdf}$  and  $\tilde{\kappa}_1 \mathcal{K}_{pdf}$  of the mimicked  $\mathcal{K}$  law (from equation (6)), given its known parameters  $L_{\mathcal{K}}$  and  $M_{\mathcal{K}}$ .
- 2) Identify these numerical values with the analytical expressions of  $\tilde{\kappa}_2 \mathcal{F}_{pdf}$ ,  $\tilde{\kappa}_3 \mathcal{F}_{pdf}$  and then  $\tilde{\kappa}_1 \mathcal{F}_{pdf}$  or  $\tilde{\kappa}_2 \mathcal{G}_{pdf}$ ,  $\tilde{\kappa}_3 \mathcal{G}_{pdf}$  and then  $\tilde{\kappa}_1 \mathcal{G}_{pdf}$  (see equations (2) and (4)), in

order to estimate respectively the parameters  $L_{\mathcal{F}}$ ,  $M_{\mathcal{F}}$  and then  $\mu_{\mathcal{F}}$  or  $L_{\mathcal{G}\mathcal{G}}$ ,  $\eta_{\mathcal{G}\mathcal{G}}$  and then  $\mu_{\mathcal{G}\mathcal{G}}$  of the mimicking laws.

In the case 2, the log-cumulant based parameter estimation of the mimicking law (here, the Fisher law or the Generalized Gamma law), likely to imitate the reference  $\mathcal{K}$  law, can be indirectly done using the three first empirical log-cumulants of the  $\mathcal{K}$  distribution through the following steps:

- 1) Compute the empirical log-cumulants  $\widehat{\kappa}_2$ ,  $\widehat{\kappa}_3$  and then  $\widehat{\kappa}_1$  of the mimicked  $\mathcal{K}$  law, given some sample values distributed according to the  $\mathcal{K}$  law, from the following formula:

$$\begin{aligned}\widehat{\kappa}_2 &= \frac{1}{N} \sum_{i=1}^N (\ln(x_i))^2 - \frac{1}{N^2} \left( \sum_{i=1}^N \ln(x_i) \right)^2 \\ \widehat{\kappa}_3 &= \frac{1}{N} \sum_{i=1}^N (\ln(x_i))^3 \\ &\quad - \frac{3}{N^2} \left( \sum_{i=1}^N \ln(x_i) \right) \left( \sum_{i=1}^N (\ln(x_i))^2 \right) \\ &\quad + \frac{2}{N^3} \left( \sum_{i=1}^N \ln(x_i) \right)^3 \\ \widehat{\kappa}_1 &= \frac{1}{N} \sum_{i=1}^N \ln(x_i)\end{aligned}\quad (14)$$

where  $x_i$  is the pixel intensity value and  $N$  is the number of available pixels.

- 2) As for step 2) of the case 1, identify these numerical values with the analytical expressions of  $\widetilde{\kappa}_2_{\mathcal{F}_{pdf}}$ ,  $\widetilde{\kappa}_3_{\mathcal{F}_{pdf}}$  and then  $\widetilde{\kappa}_1_{\mathcal{F}_{pdf}}$  or  $\widetilde{\kappa}_2_{\mathcal{G}\mathcal{G}_{pdf}}$ ,  $\widetilde{\kappa}_3_{\mathcal{G}\mathcal{G}_{pdf}}$  and then  $\widetilde{\kappa}_1_{\mathcal{G}\mathcal{G}_{pdf}}$  (see equations (2) and (4)), in order to estimate the parameters  $L_{\mathcal{F}}$ ,  $M_{\mathcal{F}}$  and then  $\mu_{\mathcal{F}}$  or  $L_{\mathcal{G}\mathcal{G}}$ ,  $\eta_{\mathcal{G}\mathcal{G}}$  and then  $\mu_{\mathcal{G}\mathcal{G}}$  of the mimicking laws.

To invert equations (2) and (4), involving especially Digamma and Polygamma functions, the algorithm described in details in [32] has been directly used for the parameter estimation of the Fisher law and easily adapted for the parameter estimation of the Generalized Gamma one.

### B. Qualitative and quantitative criteria of comparison

The evaluation of the Fisher and Generalized Gamma mimicking performances is based on qualitative and quantitative criteria, taking into account all estimated parameters simultaneously.

1) *Qualitative criterion:* For case 1, the adequation between the pdf of the mimicked  $\mathcal{K}$  law (given its known parameters) and the pdf of the mimicking Fisher or Generalized Gamma law (once its parameters have been estimated using the true log-cumulants of  $\mathcal{K}$ ) can be visually appreciated by plotting both of them on a common graph.

2) *Quantitative criteria:* For cases 1 and 2, we propose a quantitative criterion based on the Kullback-Leibler divergence (KLD) [33], that measures the dissimilarity between two distributions. We remind that the KLD between the distributions  $\mathcal{P}$  and  $\mathcal{Q}$  is defined by

$$KLD(\mathcal{P}_{pdf}(x), \mathcal{Q}_{pdf}(x)) = \int_0^{\infty} \mathcal{P}_{pdf}(x) \ln \left( \frac{\mathcal{P}_{pdf}(x)}{\mathcal{Q}_{pdf}(x)} \right) dx \quad (15)$$

This non-symmetric version allows us to give to the KLD a physical interpretation, namely a proximity measure between

an empirical distribution and a distribution (Sanov theorem, [34] [35]), whose use will be suitable for our mimicking framework. Indeed, the  $\mathcal{P}$  distribution - usually corresponding in such a non-symmetric KLD formulation to an “observed” distribution or a “true” one - refers in this paper to the mimicked one (for instance a  $\mathcal{K}$  one), while the  $\mathcal{Q}$  distribution - usually corresponding in such a non-symmetric KLD formulation to a “model” distribution that “describes, approximates or imitates” - refers in this paper to the mimicking one (i.e. here Fisher or Generalized Gamma). Therefore the main advantage of using here this non-symmetric definition is to preserve such respective roles (reference law/estimated law) and thus to keep this important and appropriate physical meaning, even if such a non-symmetric divergence does not match with a distance in a mathematical sense (which is not required for our purpose).

**For case 1**, we propose to compute a single KLD between  $\mathcal{K}$  and Fisher and a single KLD between  $\mathcal{K}$  and Generalized Gamma through the following steps:

- 1) we fix a set of parameters describing a mimicked  $\mathcal{K}$  law
- 2) we estimate the parameters of the mimicking Fisher (or Generalized Gamma) law through the log-cumulant based approach detailed in previous section for case 1
- 3) we compute the KLD between the mimicked  $\mathcal{K}$  law (given its fixed parameters) and the Fisher (or Generalized Gamma) law that has been estimated using the true log-cumulants of  $\mathcal{K}$  (computed from its known parameters).

**For case 2**, we propose to compute a mean and a variance on several KLD between  $\mathcal{K}$  and Fisher and a mean and a variance on several KLD between  $\mathcal{K}$  and Generalized Gamma through the following steps:

- 1) we fix a set of parameters describing a mimicked  $\mathcal{K}$  law and a given sample size
- 2) we generate several realizations of sample values that are distributed according to this  $\mathcal{K}$  law
- 3) for each realization, we estimate the parameters of the mimicking Fisher (or Generalized Gamma) law through the log-cumulant based approach detailed in the previous section for case 2
- 4) for each realization, we compute the KLD between the mimicked  $\mathcal{K}$  law (given its fixed parameters) and the Fisher (or Generalized Gamma) law that has been estimated using the empirical log-cumulants of  $\mathcal{K}$  (computed from sample values)
- 5) we compute the mean (for precision analysis purpose) and the variance (for robustness analysis purpose) on the different realizations of these KLD (considering either Fisher or Generalized Gamma as mimicking law).

In this framework, the words “precise” and “robust”, used to analyze the capabilities of both mimicking laws, refer to the following respective definitions: a small KLD for case 1 - or a small mean of KLD for case 2 - will be characteristic of a good precision, while a small variance of KLD for case 2 will be characteristic of a good robustness.

In addition, let us denote that it is possible that the variances of the estimators, induced by the selected MLC method, have an influence on the performances observed in terms of KLD

means and variances, that would then not depend only on the intrinsic Fisher or Generalized Gamma capacities for  $\mathcal{K}$  law mimicking. Thus, the conclusions drawn in the following about possible predominance of a given law based on precision-robustness considerations should be then cautiously interpreted. It is important to keep in mind that their validity is limited to a well-defined framework (i.e. based on empirical KLD mean and variance estimations without accounting for the variance of the MLC estimator itself). Nevertheless, it remains interesting to compare such performances because, in practice, the MLC is often the single usable method, as explained in section I. Moreover, in order to complete this outlook, the reader could also refer to [10], where the variances of different estimators, and in particular the ones induced by the MLC approach, are analytically and numerically studied in details. Such a point is however beyond the scope of this paper.

### C. Analytical formula of the Kullback-Leibler divergences between $\mathcal{K}$ / Fisher and between $\mathcal{K}$ / Generalized Gamma

The analytical expressions of the KLD between the mimicked  $\mathcal{K}$  law and the estimated Fisher law or the estimated Generalized Gamma law have the following simplified forms (see Appendices A, B and C for more details):

$$\begin{aligned} KLD(\mathcal{K}_{pdf}, \mathcal{F}_{pdf}) &= A_{\mathcal{K}/\mathcal{F}} \\ &\quad - (B_{\mathcal{K}/\mathcal{F}} + C_{\mathcal{K}/\mathcal{F}} + D_{\mathcal{K}/\mathcal{F}}) \\ KLD(\mathcal{K}_{pdf}, \mathcal{G}\mathcal{G}_{pdf}) &= A_{\mathcal{K}/\mathcal{G}\mathcal{G}} \\ &\quad - (B_{\mathcal{K}/\mathcal{G}\mathcal{G}} + C_{\mathcal{K}/\mathcal{G}\mathcal{G}} + D_{\mathcal{K}/\mathcal{G}\mathcal{G}}) \end{aligned} \quad (16)$$

with

$$\begin{aligned} A_{\mathcal{K}/\mathcal{F}} &= \int_0^{\infty} \mathcal{K}_{pdf}(x) \ln(\mathcal{K}_{pdf}(x)) dx \\ B_{\mathcal{K}/\mathcal{F}} &= \ln\left(\frac{L_{\mathcal{F}}}{M_{\mathcal{F}}\mu_{\mathcal{F}}} \frac{\Gamma(L_{\mathcal{F}}+M_{\mathcal{F}})}{\Gamma(L_{\mathcal{F}})\Gamma(M_{\mathcal{F}})}\right) \\ C_{\mathcal{K}/\mathcal{F}} &= (L_{\mathcal{F}} - 1) \left( \ln\left(\frac{L_{\mathcal{F}}}{M_{\mathcal{F}}\mu_{\mathcal{F}}}\right) + \widetilde{\kappa}_1 \mathcal{K}_{pdf} \right) \\ D_{\mathcal{K}/\mathcal{F}} &= -(L_{\mathcal{F}} + M_{\mathcal{F}}) \int_0^{\infty} \mathcal{K}_{pdf}(x) \ln\left(1 + \frac{L_{\mathcal{F}}x}{M_{\mathcal{F}}\mu_{\mathcal{F}}}\right) dx \end{aligned} \quad (17)$$

and

$$\begin{aligned} A_{\mathcal{K}/\mathcal{G}\mathcal{G}} &= \int_0^{\infty} \mathcal{K}_{pdf}(x) \ln(\mathcal{K}_{pdf}(x)) dx \\ B_{\mathcal{K}/\mathcal{G}\mathcal{G}} &= \ln\left(\frac{|\eta_{\mathcal{G}\mathcal{G}}|}{\mu_{\mathcal{G}\mathcal{G}}} \frac{L_{\mathcal{G}\mathcal{G}}^{(\frac{1}{\eta_{\mathcal{G}\mathcal{G}}})}}{\Gamma(L_{\mathcal{G}\mathcal{G}})}\right) \\ C_{\mathcal{K}/\mathcal{G}\mathcal{G}} &= (\eta_{\mathcal{G}\mathcal{G}} L_{\mathcal{G}\mathcal{G}} - 1) \left( \ln\left(\frac{L_{\mathcal{G}\mathcal{G}}^{(\frac{1}{\eta_{\mathcal{G}\mathcal{G}}})}}{\mu_{\mathcal{G}\mathcal{G}}}\right) + \widetilde{\kappa}_1 \mathcal{K}_{pdf} \right) \\ D_{\mathcal{K}/\mathcal{G}\mathcal{G}} &= -\frac{L_{\mathcal{G}\mathcal{G}}}{\mu_{\mathcal{G}\mathcal{G}}^{\eta_{\mathcal{G}\mathcal{G}}}} \int_0^{\infty} \mathcal{K}_{pdf}(x) x^{\eta_{\mathcal{G}\mathcal{G}}} dx \end{aligned} \quad (18)$$

where  $\Gamma(x)$  is the Gamma function and  $\widetilde{\kappa}_1 \mathcal{K}_{pdf}$  is the log-cumulant of order 1 of the  $\mathcal{K}$  law.

For given values of the parameters, it is possible to have a numerical evaluation of these analytical expressions using for instance Maple software or Python routines with `mpmath` library [36].

## IV. RESULTS ON DATASETS OF FISHER AND GENERALIZED GAMMA PERFORMANCES FOR $\mathcal{K}$ MIMICKING

In this section, we present the results obtained on six studied cases to compare the ability of Fisher and Generalized Gamma to mimick a reference  $\mathcal{K}$  law.

### A. Description of the sets of parameters

The six following sets of parameters have been tested:  $\mathcal{K}(1, 1, 1)$  (equal to  $\mathcal{K}_c(1, 1)$ ),  $\mathcal{K}(1, 2, 2)$  (equal to  $\mathcal{K}_c(1, 2)$ ),  $\mathcal{K}(1, 4, 4)$  (equal to  $\mathcal{K}_c(1, 4)$ ),  $\mathcal{K}(1, 1, 3)$ ,  $\mathcal{K}(1, 2, 4)$  and  $\mathcal{K}(1, 3, 5)$ .

The three first sets correspond to some special cases where  $L_{\mathcal{K}} = M_{\mathcal{K}}$  and they are thus associated to points located on the  $\mathcal{K}_c$  caustic curve (see stars on figure 1). The three last sets correspond to some general cases and they are thus associated to points located in the  $\mathcal{K}$  area (see bullets on figure 1).

The choice of the special  $\mathcal{K}_c$  caustic case for the three first sets is motivated by a general outlook larger than the one exposed in this paper; indeed, this may aim to find a potential curve in the  $\widetilde{\kappa}_2 - \widetilde{\kappa}_3$  diagram likely to delimitate “favored” domains, where the preferential use of either Fisher or Generalized Gamma distributions could be recommended. Even if this global goal is out of the focus of this paper, it is interesting to analyze here the special  $\mathcal{K}_c$  caustic case because of the following properties: firstly, the  $\mathcal{K}_c$  law corresponds to the more “balanced” one among all the  $\mathcal{K}$  laws (same weight attributed to the head and to the tail of the pdf); secondly, the  $\mathcal{K}_c$  “intermediate” location in the  $\widetilde{\kappa}_2 - \widetilde{\kappa}_3$  diagram makes it intuitively a good candidate law to lead a comparison between both Fisher and Generalized Gamma mimicking performances (indeed, above the  $\mathcal{K}_c$  curve, when  $\eta_{\mathcal{G}\mathcal{G}}$  tends towards zero, Generalized Gamma is no more analytically defined at the ordinate axis; while, below the  $\mathcal{K}_c$  curve, when  $M_{\mathcal{F}}$  tends towards the Infinite, Fisher domain is delimited by the Gamma curve).

### B. Presentation of the qualitative and quantitative results

For the case 2, 50 realizations of samples have been generated using Maple software for each set of parameters and for each sample size to compute the means and the variances of the KLD. Different sample sizes  $N_s$  have been considered between 250 and 12500 ( $N_s = 250, 500, 1000, 5000, 10000, 12500$ ).

In figure 2 (a)-(f) are presented the fittings between the pdfs of the  $\mathcal{K}$  mimicked laws and the pdfs of the Fisher or Generalized Gamma mimicking laws, that are estimated in case 1 (i.e. using the true log-cumulants of  $\mathcal{K}$ ). In figure 3 (a)-(f) are presented the evolutions in function of sample sizes  $N_s$  of the means of the KLD computed between the  $\mathcal{K}$  mimicked laws and the Fisher or Generalized Gamma mimicking laws, that are estimated in case 2 (i.e. using the empirical log-cumulants of  $\mathcal{K}$ ). On the same graphs are superimposed the constant values of the KLD computed between the  $\mathcal{K}$  mimicked laws and the Fisher or Generalized Gamma mimicking laws, that are estimated in case 1 (i.e. using the true log-cumulants of  $\mathcal{K}$ ). In figure 4 (a)-(f) are presented the evolutions in function

of sample sizes  $N_s$  of the variances of the KLD computed between the  $\mathcal{K}$  mimicked laws and the Fisher or Generalized Gamma mimicking laws, that are estimated in case 2 (i.e. using the empirical log-cumulants of  $\mathcal{K}$ ).

### C. Observations and discussion

For the six sets of parameters, the Generalized Gamma law mimicks visually the reference  $\mathcal{K}$  law slightly better than the Fisher law, although both laws provide a very good adequation (see figure 2 (a)-(f)).

This slight predominance is confirmed by the quantitative study. The constant KLD values computed between the mimicked  $\mathcal{K}$  law and the Generalized Gamma law estimated in case 1 are smaller for the six sets of parameters than those computed between the mimicked  $\mathcal{K}$  law and the Fisher law estimated in case 1 (see figure 3 (a)-(f)). The KLD means computed between the mimicked  $\mathcal{K}$  law and the Generalized Gamma or Fisher laws estimated in case 2 converge towards the constant KLD values estimated in case 1, when the sample size  $N_s$  tends towards the Infinite. For each  $N_s$  value, we observe that the KLD means computed with the Generalized Gamma laws estimated in case 2 are below those computed with the Fisher laws estimated in case 2. This demonstrates that, for the six sets of parameters, Generalized Gamma is slightly better than Fisher in precision for mimicking the  $\mathcal{K}$  law, if we restrict our point of view to a study exclusively based on empirical KLD mean and variance estimations without accounting for the variance of the MLC estimator itself. Finally, concerning the KLD variances, the observed tendencies depend on the considered set of parameters (see figure 4 (a)-(f)). For the first set ( $\mathcal{K}(1, 1, 1)$ ), the KLD variances computed with the Generalized Gamma laws estimated in case 2 are slightly below those computed with the Fisher laws estimated in case 2 for each  $N_s$  value. Nevertheless, for the last five sets, this is true only for  $N_s$  values large enough (when  $N_s$  is respectively over 1000, 5000, 500, 1000 and 5000). This demonstrates that, for these last five sets, Fisher is slightly better than Generalized Gamma in robustness for mimicking the  $\mathcal{K}$  law for small  $N_s$  values and this is the opposite for large  $N_s$  values, once again from a restricted point of view.

We summarize below some global conclusions issued from this study, aiming to compare the abilities of Fisher and Generalized Gamma to mimick a reference  $\mathcal{K}$  law (please keep in mind that these conclusions are valid for a study exclusively based on empirical KLD mean and variance estimations without accounting for the variance of the MLC estimator itself): 1) A quite good visual mimicking has been obtained with the Fisher law as well as with the Generalized Gamma law. 2) The Generalized Gamma law seems slightly better than the Fisher law in precision for all sample sizes for the six studied sets of parameters. 3) The predominance in robustness depends on the considered set of parameters and on the sample size, with the main following tendency: The Fisher law seems slightly better than the Generalized Gamma law in robustness for small sample sizes (as it can be the case sometimes in operational conditions) and this seems to be the opposite for sufficiently large sample sizes.

Thus, despite the quasi equivalent visual mimicking ability of both laws, it appears quantitatively, for the six studied sets of parameters, that it is preferable to use the Generalized Gamma law instead of the Fisher law for mimicking purpose of the  $\mathcal{K}$  law (when it is possible), if we limit our analysis to the precision aspect.

### V. MIMICKING OTHER USUAL SAR DISTRIBUTIONS: WEIBULL, BETA AND LOG-NORMAL

In this section, we present the results obtained on three studied cases to compare the ability of Fisher and Generalized Gamma to mimick some reference laws. Depending on the mimicked pdf, Fisher or Generalized Gamma may be more or less adapted.

The chosen Weibull pdf  $\mathcal{W}(1, 0.8)$ , the chosen Beta  $\mathcal{B}(1, 1.2, 3.4)$  and log-normal  $\mathcal{L}(1, 0.6)$  are indicated in the  $\widetilde{\kappa}_2 - \widetilde{\kappa}_3$  diagram of figure 1 (on the Weibull curve, in the Beta area and near the ordinate axis of log-normal respectively).

Figure 5 (a)-(b)-(c) presents the fittings between the pdfs of the Weibull, Beta and log-normal mimicked laws and the pdfs of the Fisher or Generalized Gamma mimicking laws. In table I are provided the corresponding values of the KLD numerically computed between the mimicked laws and the Fisher or Generalized Gamma laws.

Let us denote again that Weibull is a special case of Generalized Gamma distribution. It is thus perfectly mimicked by the Generalized Gamma pdf and the KLD is 0. It can also be seen on figure 5 that the Fisher pdf mimics also very well the Weibull pdf, with a small KLD value.

Concerning the Beta pdf, both Fisher and Generalized Gamma give inaccurate estimations but the approximation is better for the Generalized Gamma pdf than for Fisher. It is seen both on the figure 5 (specially for the head of the pdf) and in the KLD values (table I). Since the Fisher pdf does not cover this part of the  $\widetilde{\kappa}_2 - \widetilde{\kappa}_3$  diagram, this result was expected. It confirms the higher generality of the Generalized Gamma pdf.

When trying to mimick the Log-Normal law,  $\mathcal{L}(1, 0.6)$ , defined by one mean parameter  $\mu_{\mathcal{L}}$  and one standard deviation parameter  $\sigma_{\mathcal{L}}$ , we also got a satisfying estimation when using a Fisher law, but slightly less satisfying when using a Generalized Gamma law (through the original asymptotical approach proposed in [22] and [23] for log-normal approximation and based on a specific constrained parametrization of the Generalized Gamma), due to a too small  $\eta_{GG}$  value and a too big  $L_{GG}$  value. Note that when  $\kappa_3$  tends towards 0, the associated  $L_{GG}$  value becomes very high. In practice to avoid numerical difficulties the  $L_{GG}$  value should be limited to 60 (for such a value  $\Gamma(L) = 1.38683e + 80$ ).

### VI. FISHER AND GENERALIZED GAMMA MUTUAL MIMICKING

In this section, we present the results obtained on two studied cases to compare the ability of Fisher and Generalized Gamma to mimick each other. Indeed, since Fisher and Generalized Gamma flexibilities are due to different reasons (two head/tail weighting shape parameters for the first and one

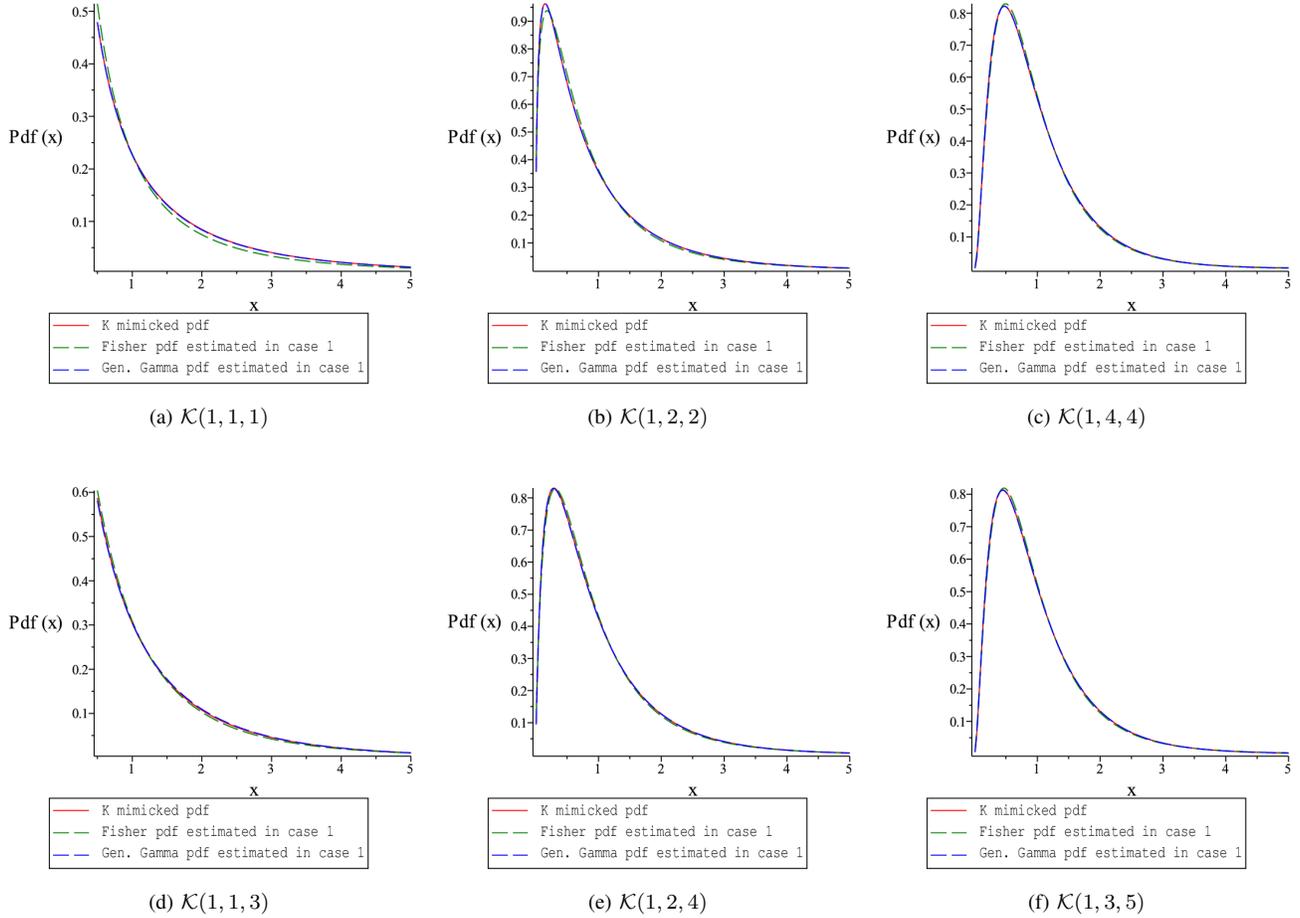


Fig. 2. Probability density functions of the  $\mathcal{K}$  mimicked laws (in red-solid) and of the mimicking laws (in green-dash for Fisher and blue-dash for Generalized Gamma) estimated in case 1 for the six sets of parameters.

TABLE I  
NUMERICAL EVALUATIONS OF KULLBACK-LEIBLER DIVERGENCES FOR THE MIMICKED-MIMICKING STUDIED CASES

	Mimicked $\mathcal{W}(1, 0.8)$	Mimicked $\mathcal{B}(1, 1.2, 3.4)$	Mimicked $\mathcal{L}(1, 0.6)$	Mimicked $\mathcal{F}(1, 0.8, 1.2)$	Mimicked $\mathcal{GG}(1, 4.5, 0.25)$
Mimicking Fisher law	0.00113	0.0608	0.00004	-	0.011942
Mimicking Generalized Gamma law	0.0	0.0165	0.00141	0.003279	-

shape parameter combined with one power parameter whose sign is either positive or negative for the second). It is thus interesting to study if they are able to mimic each other, in order to determine whether our approach and conclusions of sections IV and V can be generalized to other kinds of mimicked laws. Here also, the presented study is restricted to the situation denoted by case 1 in the previous sections III and IV.

#### A. Description of the sets of parameters

The two following sets of parameters have been tested:  $\mathcal{F}(1, 0.8, 1.2)$  and  $\mathcal{GG}(1, 4.5, 0.25)$ . Both sets are associated to points located in the Fisher area (common to Fisher and Generalized Gamma) and on the part to the left of the ordinate axis (see triangles on figure 1).

#### B. Analytical formula of the Kullback-Leibler divergences between Fisher / Generalized Gamma and Generalized Gamma / Fisher

The analytical expressions of the KLD between the mimicked Fisher law and the estimated Generalized Gamma law and between the mimicked Generalized Gamma law and the estimated Fisher law have the simplified forms respectively detailed in Appendices D and E, that represent another contribution of this paper.

#### C. Presentation of the qualitative and quantitative results

In figure 6 (a)-(b) are presented the fittings between the pdfs of the Fisher and Generalized Gamma laws mutually mimicked the one by the other and that are here also estimated in case 1 (i.e. using their true log-cumulants). In table I are provided the corresponding values of the KLD numerically computed

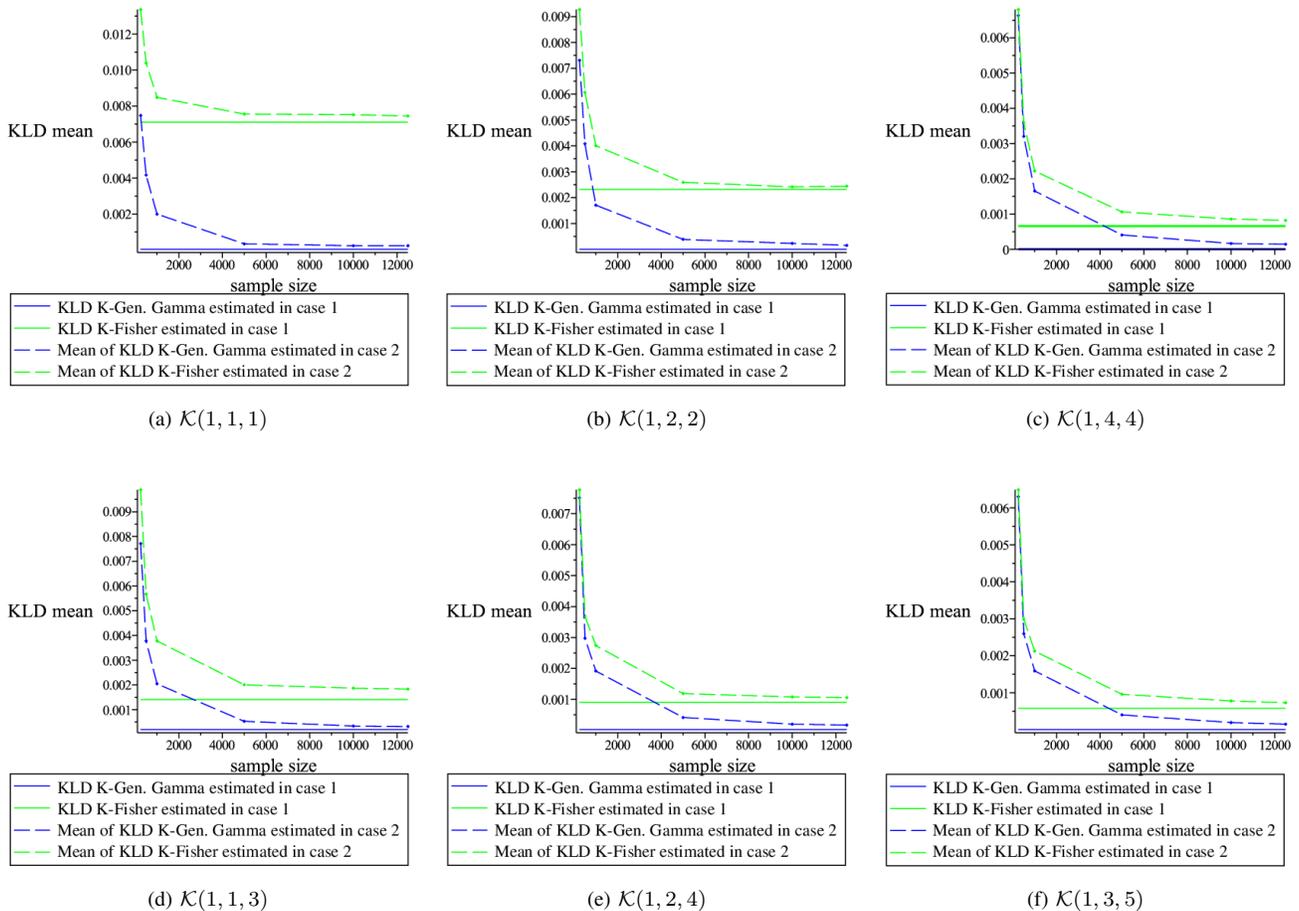


Fig. 3. Evolutions in function of sample sizes  $N_s$  of the means of the Kullback-Leibler divergences computed between the  $\mathcal{K}$  mimicked laws and the mimicking laws (in green-dash for Fisher and blue-dash for Generalized Gamma) estimated in case 2 for the six sets of parameters. The constant values of the Kullback-Leibler divergences computed between the  $\mathcal{K}$  mimicked laws and the mimicking laws (in green-solid for Fisher and blue-solid for Generalized Gamma) estimated in case 1 are superimposed on the same graphs.

between the Fisher mimicked law and the Generalized Gamma mimicking law for the first example and between the Generalized Gamma mimicked law and the Fisher mimicking law for the second example.

#### D. Observations and discussion

We can appreciate on figures 6 (a) and (b), that, for both examples, the mimicking law fits visually very well the reference mimicked law. Besides, despite that the KLD values obtained on table I for these two examples can not be directly compared, neither between them (since the used KLD are not symmetric), nor with the KLD obtained in sections IV and V (since the mimicked reference laws are not of the same kind), we can notice nevertheless that their relative order of magnitude is coarsely similar (about a factor four). We can thus conclude that, for the considered illustrative studied cases, Fisher and Generalized Gamma distributions are sufficiently flexible to correctly mimick each other. By comparison to other laws, both distributions are thus at the same time flexible, generic (covering a large area in the log-cumulant diagram) and usable in a realistic way (only three parameters that need to be estimated in practice).

## VII. RECOMMENDATIONS AND APPLICATION TO THREE REAL SAR SUBSCENES

In this section, we propose to first discuss general theoretical and practical aspects that should be taken into account when using a probability density function for real SAR data modeling. Then, we apply such recommendations on three applicative examples, to simultaneously: (i) show the very good potential of both Fisher and Generalized Gamma distributions for high-resolution data modeling; (ii) underline also the importance of cautiously choosing when possible the best suitable modeling law and when available the more appropriate samples for the given log-cumulant based estimation method; (iii) and discuss the limits of the proposed approach due to numerical difficulties.

### A. Pragmatical recommendations

We propose here to draw up pragmatical recommendations about the preferential use of a distribution in particular, even when using a flexible law for intensity data modeling. Three points are highlighted: (1) As said before, the use of MLC for parameter estimation should be verified, given the location in

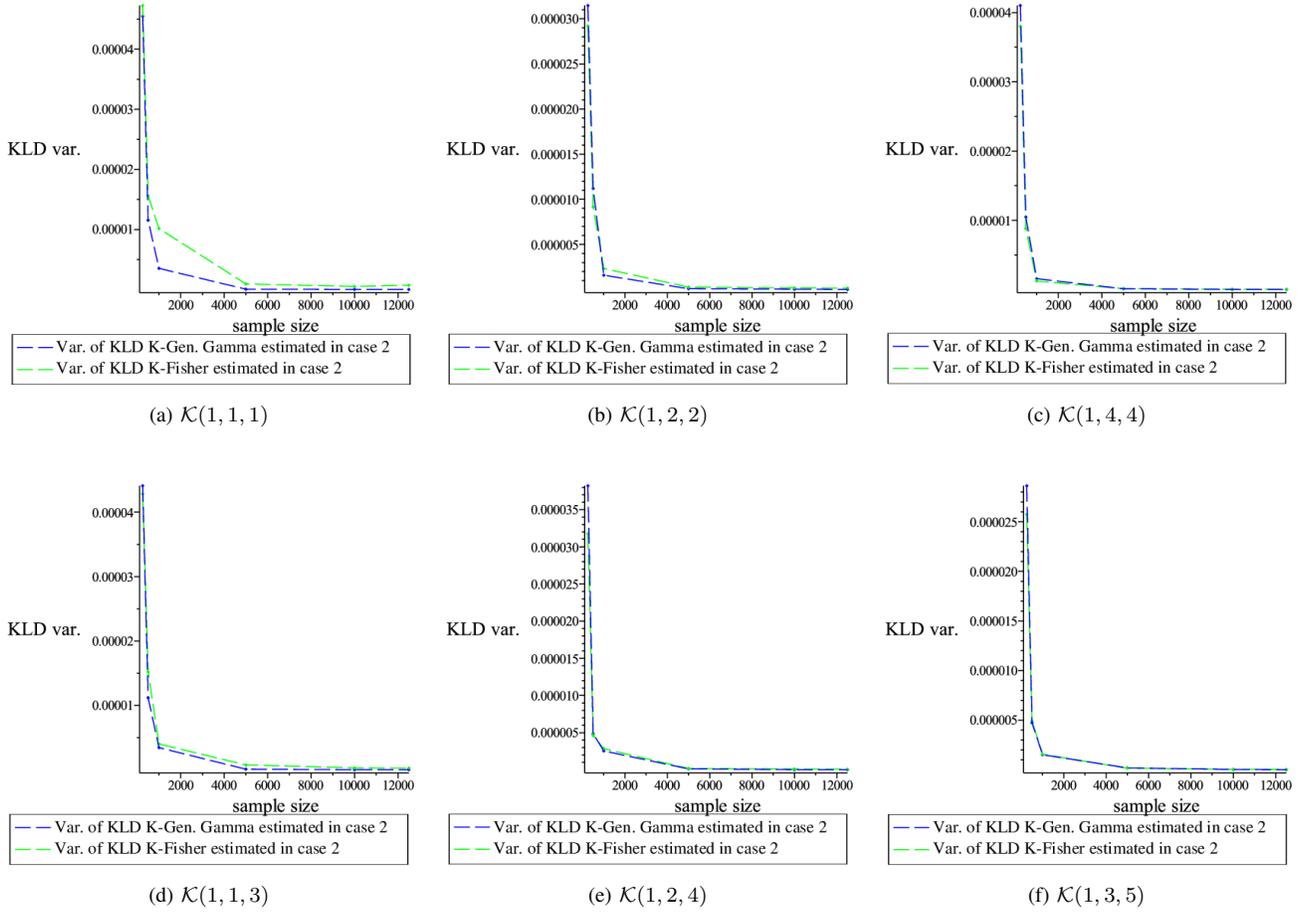


Fig. 4. Evolutions in function of sample sizes  $N_s$  of the variances of the Kullback-Leibler divergences computed between the  $\mathcal{K}$  mimicked laws and the mimicking laws (in green-dash for Fisher and blue-dash for Generalized Gamma) estimated in case 2 for the six sets of parameters.

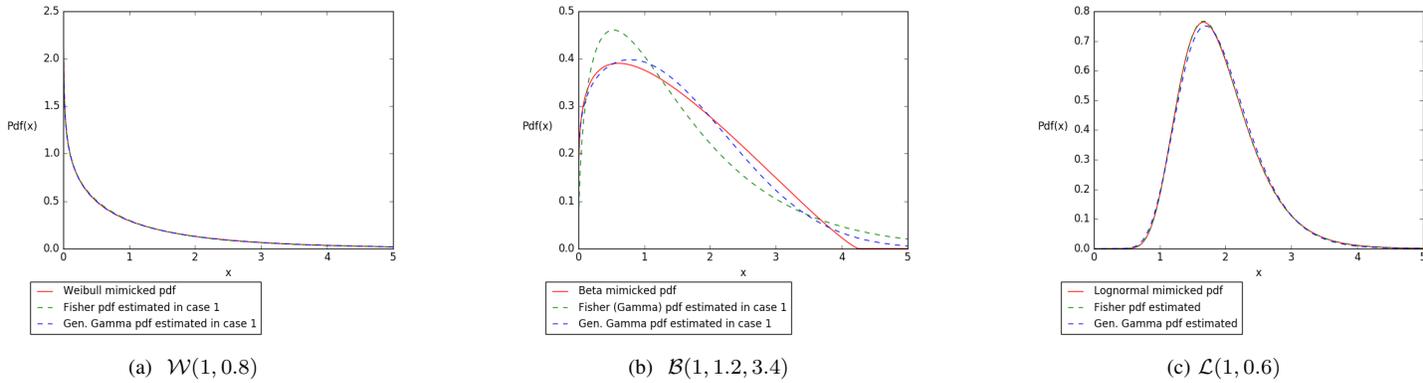


Fig. 5. Probability density functions of the Weibull (on the left), Beta (middle) and log-normal (on the right) mimicked laws and of the mimicking laws (in green-dash for Fisher and blue-dash for Generalized Gamma).

the log-cumulant diagram of the empirically estimated point  $(\widehat{\kappa}_2, \widehat{\kappa}_3)$  (usually using a local sliding window):

- for Generalized Gamma parameter estimation, equation (13) has to be checked (see [19], [28], [29] for practical cases);
- for Fisher parameter estimation, equation (11) has to be checked (see [2] for practical cases).

Moreover, the global location of the cloud of points if many estimations are computed can be used to help the decision about the best suitable law. Indeed, if the cloud sprawl is significantly going below the Gamma and/or Inverse of Gamma curves (but is staying above the curves designating the Generalized Gamma limits), the Generalized Gamma distribution seems more suitable than Fisher for modeling purpose, given its

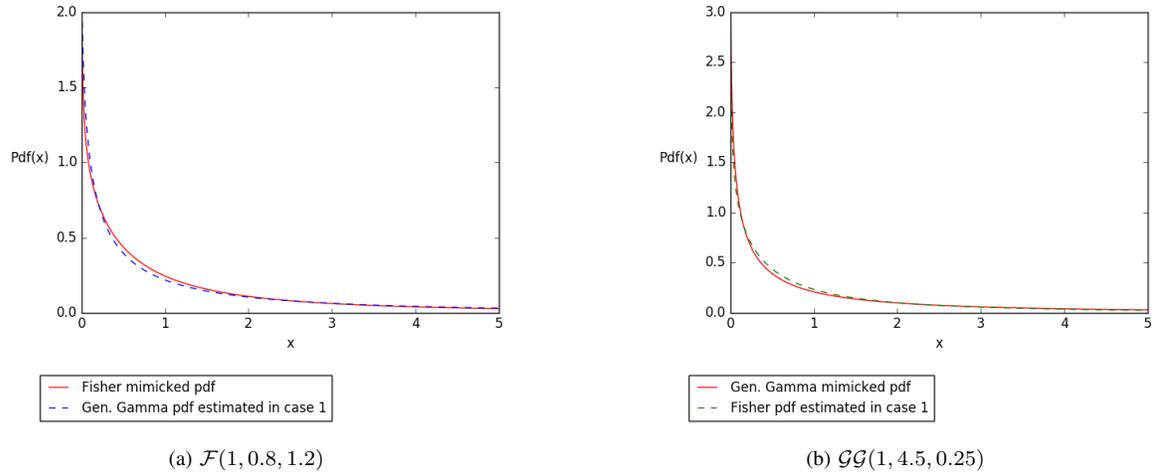


Fig. 6. Probability density functions of the mimicked laws (respectively Fisher and Generalized Gamma in red-solid) and of the mimicking laws (respectively Generalized Gamma in blue-dash and Fisher in green-dash) estimated in case 1 for both sets of parameters.

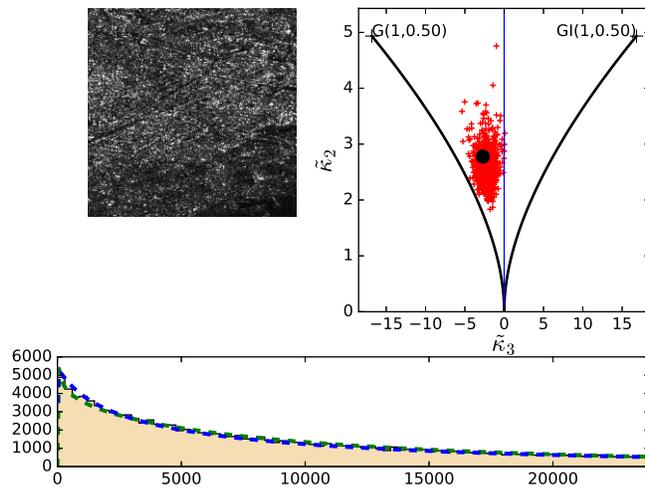


Fig. 7. (a) Upper part left: Sentinel sub-scene ©ESA (decametric resolution) acquired on Marseilles city -urban area- (acquisition date: 2014-11-08, VV polarization, descending pass, incidence angle  $39.19^\circ$ , coordinates latitude  $43^\circ 17' 17''$ N and longitude  $5^\circ 22' 47''$ E) - (b) Upper part right: Dispersion in the log-cumulant diagram of the  $(\kappa_2, \kappa_3)$  points, empirically estimated from samples. Each red point has been estimated with a 32 by 32 pixels window size. The black point has been estimated using the whole image of size 1024 by 1024 pixels. - (c) Lower part: Probability density functions of the estimated modeling laws (Generalized Gamma in blue-dash and Fisher in green-dash) superimposed on the histogram of intensity data for the whole image.

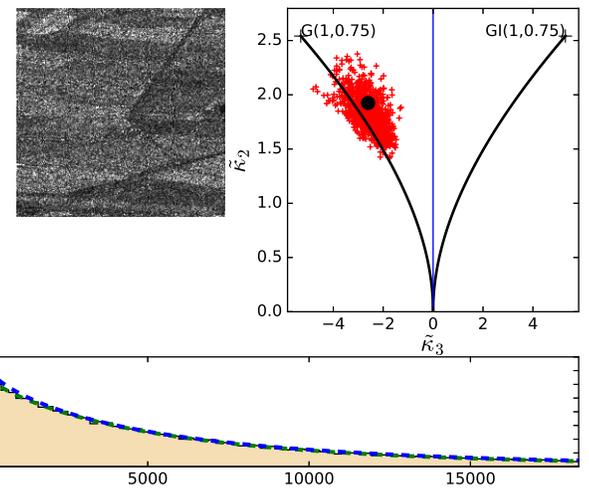


Fig. 8. (a) Upper part left: Sentinel sub-scene ©ESA (decametric resolution) acquired on Camargue area -vegetation area- (South of France) (acquisition date: 2014-11-08, VV polarization, descending pass, incidence angle  $43.83^\circ$ , coordinates latitude  $N43^\circ 28' 32''$  and longitude  $4^\circ 42' 31''$ E) - (b) Upper part right: Dispersion in the log-cumulant diagram of the  $(\kappa_2, \kappa_3)$  points, empirically estimated from samples. Each red point has been estimated with a 32 by 32 pixels window size. The black point has been estimated using the whole image of size 1024 by 1024 pixels. - (c) Lower part: Probability density functions of the estimated modeling laws (Generalized Gamma in blue-dash and Fisher in green-dash) superimposed on the histogram of intensity data for the whole image.

larger domain of existence.

(2) Similarly, intrinsic limits of distributions have also to be considered: for instance, Generalized Gamma is theoretically not defined on the ordinate axis ( $\eta_{GG}$  sign changing when crossing it), so it should be checked that the selected  $(\widehat{\kappa}_2, \widehat{\kappa}_3)$  point used for pdf parameter retrieval is not situated either along this axis or too close to it. The global location of the cloud of points can be here also used to help the law choice in a complementary way. In particular, if the cloud sprawl

is significantly crossing the ordinate axis and is presenting numerous points very close to it, the Fisher distribution seems in this case more suitable than Generalized Gamma, given the discontinuity of this last one.

(3) Eventually, numerical instabilities that can occur during Polygamma inversion involved in parameter estimation process or numerical precision limits that can be reached when evaluating huge quantities (in particular  $\Gamma(x)$  for large  $x$  values)

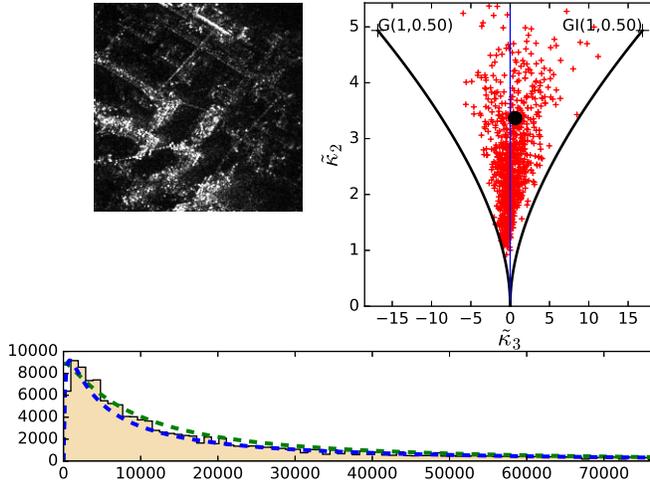


Fig. 9. (a) Upper part left: TerraSAR-X sub-scene ©DLR e. V. (2008), Distribution AIRBUS DS/Infoterra GmbH (metric resolution, SpotLight mode, area: India, Port of Visakhapatnam, acquisition date 2008-10-12, polarization VV, ascending pass, incidence angle  $41^\circ 98'$ , coordinates latitude  $27^\circ 24' 00''$  N, and longitude  $91^\circ 50' 31''$  E) - (b) Upper part right: Dispersion in the log-cumulant diagram of the  $(\widehat{\kappa}_2, \widehat{\kappa}_3)$  points, empirically estimated from samples. Each red point has been estimated with a 32 by 32 pixels window size. The black point has been estimated using the whole image of size 1024 by 1024 pixels. - (c) Lower part: Probability density functions of the estimated modeling laws (Generalized Gamma in blue-dash and Fisher in green-dash) superimposed on the histogram of intensity data for the whole image.

should be kept in mind in some critical areas:

- around the ordinate axis, where  $L_{GG}$  tends towards the Infinite and  $\eta_{GG}$  tends towards zero;
- and around the Gamma and Inverse of Gamma curves, where  $M_{\mathcal{F}}$  tends towards the Infinite.

Indeed, when some parameters are tending toward extreme values, algorithm speed often significantly decreases (because of a clearly larger number of iterations required before convergence) and sometimes the estimation can even fail (because of numerical divergence or impossibility to provide a correct numerical evaluation). Thus, selecting suitable laws in such critical cases is an aspect that should not be neglected to prevent from important errors.

Let us finally underline that, when the computation time has to be very short, an interesting alternative is the use, when it is possible, of closed-form estimators, as the one proposed in [19] for the Generalized Gamma law and based on second-kind cumulants, for which a second-order approximation of the Polygamma function allows to avoid iterative process.

### B. Application on real SAR data

Even if a complete empirical study on real SAR data is beyond the scope of this paper, we propose here to illustrate on two applicative examples (see figures 7 and 9) the potential of Fisher and Generalized Gamma laws for intensity data modeling and also to analyze in details on these examples the aspects introduced in previous section VII-A.

Figures 7, 8 and 9 present three real SAR examples respectively acquired with Sentinel sensor (decametric resolution) - urban and vegetated areas- and with Terrasar-X sensor (metric resolution). The heterogeneity level appears more important on third sub-scene than on the first and second one, in particular because of the presence of bright targets whose visibility becomes possible in high-resolution data (leading to rougher speckle and heavy tailed distributions), while isolating such targets remains impossible with lower resolution (making SAR extract to appear less heterogeneous).

Such different behaviors are depicted by the location in the log-cumulant diagram of the corresponding  $(\widehat{\kappa}_2, \widehat{\kappa}_3)$  point clouds. In the first case, the cloud is mainly located above the Gamma curve in the Fisher area (common to Fisher and Generalized Gamma) and only on the part to the left of the ordinate axis. In the second case for vegetation area, the point cloud is centered near the Gamma curve corresponding to homogeneous regions. In the opposite, in the third case, the cloud of points extends everywhere above the Gamma and Inverse of Gamma curves and is largely crossing the ordinate axis. Thus, a first fast analysis, based on the global cloud point location, would orientate us to preferentially use a Fisher law for modeling the third subscene, and indifferently a Fisher law or a Generalized Gamma law for modeling the first and second one, to prevent from any estimation difficulties, as previously explained.

As can be seen on the figures 7 and 8 with the positioning of the points, log-normal and Beta distributions are not adapted for the SAR Sentinel data. For the TerraSAR-X data with higher resolution on figure 9, some points could be explained with a log-normal distribution. As already discussed, in this case, the Fisher pdf is able to mimick efficiently the log-normal pdf.

By considering for the three cases the black  $(\widehat{\kappa}_2, \widehat{\kappa}_3)$  point, the parameters of the modeling Fisher and Generalized Gamma laws can be successfully retrieved based on the log-cumulant method. The obtained estimated distributions have been superimposed on the histograms in intensity for the three sub-scenes. A very good data fitting is observed on the results for the three examples with both laws, which shows their high flexibility.

Nevertheless, despite this noteworthy modeling ability, let us point out some correlated limits of the approach in order to raise awareness of the reader to difficulties that can be encountered for some practical cases. The third example illustrates particularly well the general recommendations discussed in section VII-A, as follows:

- First, we remind that Generalized Gamma parameter estimation has been done using the black  $(\widehat{\kappa}_2, \widehat{\kappa}_3)$  point. For this example, this point is not located “too” close to the ordinate axis, where Generalized Gamma is theoretically no defined. Thus the selection of such a point allows us to perform here the inversion successfully. The obtained estimated values of  $L_{GG}$  (tending towards the Infinite) and  $\eta_{GG}$  (tending towards zero) are respectively equal to 88.95 and -0.06. However, the inversion could have failed with another point located closer to the ordinate axis (like it is the case for many of the empirical points belonging to the

cloud on this example). We have thus to attentively retain an appropriate point and, if it is not possible, to select another more suitable modeling law (here the Fisher law) to try to handle such configurations.

- Second, given the huge numerical values involved in the inversion process for such sensitive situations (considering in particular large  $L_{GG}$  values and the corresponding computations of  $\Gamma(L_{GG})$ ), it is important to notice that the reached numerical precision will depend on the languages and softwares used for parameter estimation, and might be an imprecision factor. For instance, employing double precision C language will allow us to precisely deal only with cases where  $L_{GG}$  is strictly inferior to 20. Here, where the estimated value of  $L_{GG}$  is equal to 88.95, parameter retrieval has been done using Python routines with `mpmath` library [36].

### VIII. CONCLUSION AND FUTURE WORK

In this paper, we have compared the ability of two generic distributions, the Fisher distribution and the Generalized Gamma distribution, to correctly imitate a reference one, chosen as the  $\mathcal{K}$  law or other SAR usual laws. While the flexibility of the Fisher law relies on its two shape parameters, the Generalized Gamma law benefits especially from its power parameter: this offers to both of them an important mimicking potential, as visually confirmed by the good adequations observed on multiple examples. The proposed study for  $\mathcal{K}$  law mimicking, based on the computation of Kullback-Leibler divergences, has quantitatively demonstrated that the Generalized Gamma law is slightly better than the Fisher law in precision for the six studied sets of parameters and for all sample sizes, while the predominance in robustness depends on the considered set of parameters and on the sample size, (such conclusions being valid only from a restricted point of view). It has thus appeared as slightly preferable to use the more powerful Generalized Gamma law instead of the Fisher law, for  $\mathcal{K}$  mimicking purpose, when it was possible. The proposed study for other usual SAR laws mimicking has shown complementarities between both laws but again with a higher genericity of Generalized Gamma pdf. The recommendation and application section, based on theoretical and practical considerations applied to real SAR data, has also allowed to highlight critical situations, where even generic distributions can present some fitting limitations and where the use of given suitable modeling laws and appropriate and ingenious parameter estimation methods is very important. On the one hand, the very good potential of both Fisher and Generalized Gamma distributions for intensity data modeling has been proved on three various illustrative SAR subscenes. On the other hand, it has been shown, that independently on their intrinsic capacities, in some sensitive cases, we can be led to use preferentially a given modeling law for some practical aspects. This includes especially numerical precision limitations (in particular, it has appeared that Fisher could be more suitable than Generalized Gamma, in some cases where the  $(\tilde{\kappa}_2, \tilde{\kappa}_3)$  point used for parameter estimation is too close to the ordinate axis).

In future work, an experimental study on HR or VHR real SAR images is essential to consolidate our conclusions about Fisher and Generalized Gamma modeling power.

### APPENDIX A

#### KULLBACK-LEIBLER DIVERGENCE BETWEEN A REFERENCE LAW AND THE FISHER LAW

According to equation (15), the KLD divergence between a reference law and the Fisher law is given below, where  $\mathcal{R}_{pdf}$  is the reference pdf.

$$\begin{aligned} KLD(\mathcal{R}_{pdf}, \mathcal{F}_{pdf}) &= \int_0^{\infty} \mathcal{R}_{pdf}(x) \ln \left( \frac{\mathcal{R}_{pdf}(x)}{\mathcal{F}_{pdf}(x)} \right) dx \\ &= \underbrace{\int_0^{\infty} \mathcal{R}_{pdf}(x) \ln(\mathcal{R}_{pdf}(x)) dx}_{A_{\mathcal{R}/\mathcal{F}}} - \underbrace{\int_0^{\infty} \mathcal{R}_{pdf}(x) \ln(\mathcal{F}_{pdf}(x)) dx}_{TF_2}. \end{aligned}$$

Using equation 1), the term  $TF_2$  can be written as:

$$\begin{aligned} TF_2 &= \underbrace{\ln \left( \frac{L_{\mathcal{F}}}{M_{\mathcal{F}}\mu_{\mathcal{F}}} \frac{\Gamma(L_{\mathcal{F}} + M_{\mathcal{F}})}{\Gamma(L_{\mathcal{F}})\Gamma(M_{\mathcal{F}})} \right)}_{B_{\mathcal{R}/\mathcal{F}}} \\ &+ \underbrace{\int_0^{\infty} \mathcal{R}_{pdf}(x) \ln \left( \left( \frac{L_{\mathcal{F}}x}{M_{\mathcal{F}}\mu_{\mathcal{F}}} \right)^{L_{\mathcal{F}}-1} \right) dx}_{TF_3} \\ &+ \underbrace{\int_0^{\infty} \mathcal{R}_{pdf}(x) \ln \left( \frac{1}{\left( 1 + \frac{L_{\mathcal{F}}x}{M_{\mathcal{F}}\mu_{\mathcal{F}}} \right)^{L_{\mathcal{F}}+M_{\mathcal{F}}}} \right) dx}_{TF_4}. \end{aligned}$$

By developing and simplifying the terms  $TF_3$  and  $TF_4$ , we obtain that:

$$\begin{aligned} TF_3 &= (L_{\mathcal{F}} - 1) \int_0^{\infty} \mathcal{R}_{pdf}(x) \ln \left( \frac{L_{\mathcal{F}}x}{M_{\mathcal{F}}\mu_{\mathcal{F}}} \right) dx \\ &= (L_{\mathcal{F}} - 1) \left( \ln \left( \frac{L_{\mathcal{F}}}{M_{\mathcal{F}}\mu_{\mathcal{F}}} \right) + \underbrace{\int_0^{\infty} \mathcal{R}_{pdf}(x) \ln(x) dx}_{\tilde{\kappa}_1 \mathcal{R}_{pdf}} \right) \\ &= \underbrace{(L_{\mathcal{F}} - 1) \left( \ln \left( \frac{L_{\mathcal{F}}}{M_{\mathcal{F}}\mu_{\mathcal{F}}} \right) + \tilde{\kappa}_1 \mathcal{R}_{pdf} \right)}_{C_{\mathcal{R}/\mathcal{F}}} \end{aligned}$$

where  $\tilde{\kappa}_1 \mathcal{R}_{pdf}$  is the first log-cumulant of the reference law, and

$$TF_4 = \underbrace{-(L_{\mathcal{F}} + M_{\mathcal{F}}) \int_0^{\infty} \mathcal{R}_{pdf}(x) \ln \left( 1 + \frac{L_{\mathcal{F}}x}{M_{\mathcal{F}}\mu_{\mathcal{F}}} \right) dx}_{D_{\mathcal{R}/\mathcal{F}}}.$$

In summary, the KLD divergence using the Fisher distribution is defined by:

$$KLD(\mathcal{R}_{pdf}, \mathcal{F}_{pdf}) = \frac{A_{\mathcal{R}/\mathcal{F}}}{-(B_{\mathcal{R}/\mathcal{F}} + C_{\mathcal{R}/\mathcal{F}} + D_{\mathcal{R}/\mathcal{F}})}.$$

## APPENDIX B

### KULLBACK-LEIBLER DIVERGENCE BETWEEN A REFERENCE LAW AND THE GENERALIZED GAMMA LAW

According to equation (15), the KLD divergence between a reference law and the Generalized Gamma law is given below, where  $\mathcal{R}_{pdf}$  is the reference pdf.

$$\begin{aligned} KLD(\mathcal{R}_{pdf}, \mathcal{GG}_{pdf}) &= \int_0^{\infty} \mathcal{R}_{pdf}(x) \ln \left( \frac{\mathcal{R}_{pdf}(x)}{\mathcal{GG}_{pdf}(x)} \right) dx \\ &= \underbrace{\int_0^{\infty} \mathcal{R}_{pdf}(x) \ln(\mathcal{R}_{pdf}(x)) dx}_{A_{\mathcal{R}/\mathcal{GG}}} - \underbrace{\int_0^{\infty} \mathcal{R}_{pdf}(x) \ln(\mathcal{GG}_{pdf}(x)) dx}_{TGG_2}. \end{aligned}$$

Using equation 3), the term  $TGG_2$  can be written as:

$$\begin{aligned} TGG_2 &= \underbrace{\ln \left( \frac{|\eta_{GG}| L_{GG}^{(\frac{1}{\eta_{GG}})}}{\mu_{GG} \Gamma(L_{GG})} \right)}_{B_{\mathcal{R}/\mathcal{GG}}} \\ &+ \underbrace{\int_0^{\infty} \mathcal{R}_{pdf}(x) \ln \left( \left( \frac{L_{GG}^{(\frac{1}{\eta_{GG}})} x}{\mu_{GG}} \right)^{(\eta_{GG} L_{GG} - 1)} \right) dx}_{TGG_3} \\ &+ \underbrace{\int_0^{\infty} \mathcal{R}_{pdf}(x) \ln \left( e^{-\left( \frac{L_{GG}^{(\frac{1}{\eta_{GG}})} x}{\mu_{GG}} \right)^{\eta_{GG}}} \right) dx}_{TGG_4}. \end{aligned}$$

By developing and simplifying the terms  $TGG_3$  and  $TGG_4$ , we obtain that:

$$\begin{aligned} TGG_3 &= (\eta_{GG} L_{GG} - 1) \int_0^{\infty} \mathcal{R}_{pdf}(x) \ln \left( \frac{L_{GG}^{(\frac{1}{\eta_{GG}})} x}{\mu_{GG}} \right) dx \\ &= (\eta_{GG} L_{GG} - 1) \left( \ln \left( \frac{L_{GG}^{(\frac{1}{\eta_{GG}})}}{\mu_{GG}} \right) + \underbrace{\int_0^{\infty} \mathcal{R}_{pdf}(x) \ln(x) dx}_{\widetilde{\kappa}_1 \mathcal{R}_{pdf}} \right) \\ &= (\eta_{GG} L_{GG} - 1) \underbrace{\left( \ln \left( \frac{L_{GG}^{(\frac{1}{\eta_{GG}})}}{\mu_{GG}} \right) + \widetilde{\kappa}_1 \mathcal{R}_{pdf} \right)}_{C_{\mathcal{R}/\mathcal{GG}}} \end{aligned}$$

where  $\widetilde{\kappa}_1 \mathcal{R}_{pdf}$  is the first log-cumulant of the reference law, and

$$TGG_4 = \underbrace{-\frac{L_{GG}}{\mu_{GG}^{\eta_{GG}}} \int_0^{\infty} \mathcal{R}_{pdf}(x) x^{\eta_{GG}} dx}_{D_{\mathcal{R}/\mathcal{GG}}}$$

where the term  $\int_0^{\infty} \mathcal{R}_{pdf}(x) x^{\eta_{GG}} dx$  corresponds to the classical moment of order  $\eta_{GG}$  of the reference law.

In summary, the KLD divergence using the Generalized Gamma distribution is defined by:

$$KLD(\mathcal{R}_{pdf}, \mathcal{GG}_{pdf}) = \frac{A_{\mathcal{R}/\mathcal{GG}}}{-(B_{\mathcal{R}/\mathcal{GG}} + C_{\mathcal{R}/\mathcal{GG}} + D_{\mathcal{R}/\mathcal{GG}})}.$$

## APPENDIX C

### PROOFS OF EQUATIONS (16), (17) AND (18)

In this section, the reference law, denoted by  $\mathcal{R}_{pdf}$  in appendices A and B, is replaced by the  $\mathcal{K}$  law.

The KLD divergences between the  $\mathcal{K}$  law and the Fisher or Generalized Gamma law are thus defined by:

$$\begin{aligned} KLD(\mathcal{K}_{pdf}, \mathcal{F}_{pdf}) &= \frac{A_{\mathcal{K}/\mathcal{F}}}{-(B_{\mathcal{K}/\mathcal{F}} + C_{\mathcal{K}/\mathcal{F}} + D_{\mathcal{K}/\mathcal{F}})} \\ KLD(\mathcal{K}_{pdf}, \mathcal{GG}_{pdf}) &= \frac{A_{\mathcal{K}/\mathcal{GG}}}{-(B_{\mathcal{K}/\mathcal{GG}} + C_{\mathcal{K}/\mathcal{GG}} + D_{\mathcal{K}/\mathcal{GG}})} \end{aligned}$$

with

$$\begin{aligned} A_{\mathcal{K}/\mathcal{F}} &= \int_0^{\infty} \mathcal{K}_{pdf}(x) \ln(\mathcal{K}_{pdf}(x)) dx \\ B_{\mathcal{K}/\mathcal{F}} &= \ln \left( \frac{L_{\mathcal{F}}}{M_{\mathcal{F}} \mu_{\mathcal{F}}} \frac{\Gamma(L_{\mathcal{F}} + M_{\mathcal{F}})}{\Gamma(L_{\mathcal{F}}) \Gamma(M_{\mathcal{F}})} \right) \\ C_{\mathcal{K}/\mathcal{F}} &= (L_{\mathcal{F}} - 1) \left( \ln \left( \frac{L_{\mathcal{F}}}{M_{\mathcal{F}} \mu_{\mathcal{F}}} \right) + \widetilde{\kappa}_1 \mathcal{K}_{pdf} \right) \\ D_{\mathcal{K}/\mathcal{F}} &= -(L_{\mathcal{F}} + M_{\mathcal{F}}) \int_0^{\infty} \mathcal{K}_{pdf}(x) \ln \left( 1 + \frac{L_{\mathcal{F}} x}{M_{\mathcal{F}} \mu_{\mathcal{F}}} \right) dx \end{aligned}$$

and

$$\begin{aligned} A_{\mathcal{K}/\mathcal{GG}} &= \int_0^{\infty} \mathcal{K}_{pdf}(x) \ln(\mathcal{K}_{pdf}(x)) dx \\ B_{\mathcal{K}/\mathcal{GG}} &= \ln \left( \frac{|\eta_{GG}| L_{GG}^{(\frac{1}{\eta_{GG}})}}{\mu_{GG} \Gamma(L_{GG})} \right) \\ C_{\mathcal{K}/\mathcal{GG}} &= (\eta_{GG} L_{GG} - 1) \left( \ln \left( \frac{L_{GG}^{(\frac{1}{\eta_{GG}})}}{\mu_{GG}} \right) + \widetilde{\kappa}_1 \mathcal{K}_{pdf} \right) \\ D_{\mathcal{K}/\mathcal{GG}} &= -\frac{L_{GG}}{\mu_{GG}^{\eta_{GG}}} \int_0^{\infty} \mathcal{K}_{pdf}(x) x^{\eta_{GG}} dx \end{aligned}$$

where  $\Gamma(x)$  is the Gamma function and  $\widetilde{\kappa}_1 \mathcal{K}_{pdf}$  is the log-cumulant of order 1 of the  $\mathcal{K}$  law.

Moreover, in  $D_{\mathcal{K}/\mathcal{GG}}$ , the term  $\int_0^{\infty} \mathcal{K}_{pdf}(x) x^{\eta_{GG}} dx$  corresponds to the classical moment of order  $\eta_{GG}$  of the  $\mathcal{K}$  law and can be expressed as [10]:

$$\begin{aligned} &\int_0^{\infty} \mathcal{K}_{pdf}(x) x^{\eta_{GG}} dx \\ &= \mu_{\mathcal{K}}^{\eta_{GG}} \frac{\Gamma(L_{\mathcal{K}} + \eta_{GG})}{L_{\mathcal{K}}^{\eta_{GG}} \Gamma(L_{\mathcal{K}})} \frac{\Gamma(M_{\mathcal{K}} + \eta_{GG})}{M_{\mathcal{K}}^{\eta_{GG}} \Gamma(M_{\mathcal{K}})} \end{aligned}$$

Let us notice that the KLD divergences between the  $\mathcal{K}_c$  caustic law and the Fisher or Generalized Gamma law can be directly deduced from equations (16), (17) and (18) by replacing  $\mathcal{K}$  by  $\mathcal{K}_c$ .

#### APPENDIX D

##### KLD DIVERGENCE BETWEEN THE FISHER LAW AND THE GENERALIZED GAMMA LAW

Using appendix B and considering the Fisher law as the reference law, we can deduce that:

$$KLD(\mathcal{F}_{pdf}, \mathcal{G}\mathcal{G}_{pdf}) = A_{\mathcal{F}/\mathcal{G}\mathcal{G}} - (B_{\mathcal{F}/\mathcal{G}\mathcal{G}} + C_{\mathcal{F}/\mathcal{G}\mathcal{G}} + D_{\mathcal{F}/\mathcal{G}\mathcal{G}})$$

with

$$\begin{aligned} A_{\mathcal{F}/\mathcal{G}\mathcal{G}} &= \int_0^{\infty} \mathcal{F}_{pdf}(x) \ln(\mathcal{F}_{pdf}(x)) dx \\ B_{\mathcal{F}/\mathcal{G}\mathcal{G}} &= \ln\left(\frac{|\eta_{\mathcal{G}\mathcal{G}}| L_{\mathcal{G}\mathcal{G}}}{\mu_{\mathcal{G}\mathcal{G}} \Gamma(L_{\mathcal{G}\mathcal{G}})}\right) \\ C_{\mathcal{F}/\mathcal{G}\mathcal{G}} &= (\eta_{\mathcal{G}\mathcal{G}} L_{\mathcal{G}\mathcal{G}} - 1) \left( \ln\left(\frac{L_{\mathcal{G}\mathcal{G}}}{\mu_{\mathcal{G}\mathcal{G}}}\right) + \widetilde{\kappa}_{1\mathcal{F}_{pdf}} \right) \\ D_{\mathcal{F}/\mathcal{G}\mathcal{G}} &= -\frac{L_{\mathcal{G}\mathcal{G}}}{\mu_{\mathcal{G}\mathcal{G}}^{\eta_{\mathcal{G}\mathcal{G}}}} \int_0^{\infty} \mathcal{F}_{pdf}(x) x^{\eta_{\mathcal{G}\mathcal{G}}} dx \end{aligned}$$

where  $\Gamma(x)$  is the Gamma function and  $\widetilde{\kappa}_{1\mathcal{F}_{pdf}}$  is the log-cumulant of order 1 of the Fisher law.

Moreover, in  $D_{\mathcal{F}/\mathcal{G}\mathcal{G}}$ , the term  $\int_0^{\infty} \mathcal{F}_{pdf}(x) x^{\eta_{\mathcal{G}\mathcal{G}}} dx$  corresponds to the classical moment of order  $\eta_{\mathcal{G}\mathcal{G}}$  of the Fisher law and can be expressed as [10]:

$$\begin{aligned} &\int_0^{\infty} \mathcal{F}_{pdf}(x) x^{\eta_{\mathcal{G}\mathcal{G}}} dx \\ &= \mu_{\mathcal{F}}^{\eta_{\mathcal{G}\mathcal{G}}} \frac{\Gamma(L_{\mathcal{F}} + \eta_{\mathcal{G}\mathcal{G}})}{L_{\mathcal{F}}^{\eta_{\mathcal{G}\mathcal{G}}} \Gamma(L_{\mathcal{F}})} \frac{\Gamma(M_{\mathcal{F}} - \eta_{\mathcal{G}\mathcal{G}})}{M_{\mathcal{F}}^{-\eta_{\mathcal{G}\mathcal{G}}} \Gamma(M_{\mathcal{F}})} \end{aligned}$$

Besides, the first term  $A_{\mathcal{F}/\mathcal{G}\mathcal{G}}$  can be developed as:

$$\begin{aligned} A_{\mathcal{F}/\mathcal{G}\mathcal{G}} &= \ln\left(\frac{L_{\mathcal{F}}}{M_{\mathcal{F}} \mu_{\mathcal{F}}} \frac{\Gamma(L_{\mathcal{F}} + M_{\mathcal{F}})}{\Gamma(L_{\mathcal{F}}) \Gamma(M_{\mathcal{F}})}\right) \\ &+ (L_{\mathcal{F}} - 1) \left( \ln\left(\frac{L_{\mathcal{F}}}{M_{\mathcal{F}} \mu_{\mathcal{F}}}\right) + \widetilde{\kappa}_{1\mathcal{F}_{pdf}} \right) \\ &- \int_0^{\infty} \mathcal{F}_{pdf}(x) \ln\left(1 + \frac{L_{\mathcal{F}} x}{M_{\mathcal{F}} \mu_{\mathcal{F}}}\right) dx \end{aligned}$$

Finally, using that (see [37], page 558, formula 14):

$$\begin{aligned} &\int_0^{\infty} \mathcal{F}_{pdf}(x) \log\left(1 + \frac{L_{\mathcal{F}} x}{M_{\mathcal{F}} \mu_{\mathcal{F}}}\right) dx \\ &= (L_{\mathcal{F}} + M_{\mathcal{F}}) (\Psi(L_{\mathcal{F}} + M_{\mathcal{F}}) - \Psi(M_{\mathcal{F}})) \end{aligned}$$

we can finally deduce that, after simplifications:

$$\begin{aligned} A_{\mathcal{F}/\mathcal{G}\mathcal{G}} &= \log\left(\frac{L_{\mathcal{F}}}{M_{\mathcal{F}} \mu_{\mathcal{F}}}\right) + \log\left(\frac{\Gamma(L_{\mathcal{F}} + M_{\mathcal{F}})}{\Gamma(L_{\mathcal{F}}) \Gamma(M_{\mathcal{F}})}\right) \\ &+ (L_{\mathcal{F}} - 1) \Psi(L_{\mathcal{F}}) + (M_{\mathcal{F}} + 1) \Psi(M_{\mathcal{F}}) \\ &- (L_{\mathcal{F}} + M_{\mathcal{F}}) \Psi(L_{\mathcal{F}} + M_{\mathcal{F}}) \end{aligned}$$

#### APPENDIX E

##### KLD DIVERGENCE BETWEEN THE GENERALIZED GAMMA LAW AND THE FISHER LAW

Using appendice A and considering the Generalized Gamma law as the reference law, we can deduce that:

$$KLD(\mathcal{G}\mathcal{G}_{pdf}, \mathcal{F}_{pdf}) = A_{\mathcal{G}\mathcal{G}/\mathcal{F}} - (B_{\mathcal{G}\mathcal{G}/\mathcal{F}} + C_{\mathcal{G}\mathcal{G}/\mathcal{F}} + D_{\mathcal{G}\mathcal{G}/\mathcal{F}})$$

with

$$\begin{aligned} A_{\mathcal{G}\mathcal{G}/\mathcal{F}} &= \int_0^{\infty} \mathcal{G}\mathcal{G}_{pdf}(x) \ln(\mathcal{G}\mathcal{G}_{pdf}(x)) dx \\ B_{\mathcal{G}\mathcal{G}/\mathcal{F}} &= \ln\left(\frac{L_{\mathcal{F}}}{M_{\mathcal{F}} \mu_{\mathcal{F}}} \frac{\Gamma(L_{\mathcal{F}} + M_{\mathcal{F}})}{\Gamma(L_{\mathcal{F}}) \Gamma(M_{\mathcal{F}})}\right) \\ C_{\mathcal{G}\mathcal{G}/\mathcal{F}} &= (L_{\mathcal{F}} - 1) \left( \ln\left(\frac{L_{\mathcal{F}}}{M_{\mathcal{F}} \mu_{\mathcal{F}}}\right) + \widetilde{\kappa}_{1\mathcal{G}\mathcal{G}_{pdf}} \right) \\ D_{\mathcal{G}\mathcal{G}/\mathcal{F}} &= -(L_{\mathcal{F}} + M_{\mathcal{F}}) \int_0^{\infty} \mathcal{G}\mathcal{G}_{pdf}(x) \ln\left(1 + \frac{L_{\mathcal{F}} x}{M_{\mathcal{F}} \mu_{\mathcal{F}}}\right) dx \end{aligned}$$

where  $\Gamma(x)$  is the Gamma function and  $\widetilde{\kappa}_{1\mathcal{G}\mathcal{G}_{pdf}}$  is the log-cumulant of order 1 of the Generalized Gamma law.

Moreover, the first term  $A_{\mathcal{G}\mathcal{G}/\mathcal{F}}$  can be developed as:

$$\begin{aligned} A_{\mathcal{G}\mathcal{G}/\mathcal{F}} &= \ln\left(\frac{|\eta_{\mathcal{G}\mathcal{G}}| L_{\mathcal{G}\mathcal{G}}}{\mu_{\mathcal{G}\mathcal{G}} \Gamma(L_{\mathcal{G}\mathcal{G}})}\right) \\ &+ (\eta_{\mathcal{G}\mathcal{G}} L_{\mathcal{G}\mathcal{G}} - 1) \left( \ln\left(\frac{L_{\mathcal{G}\mathcal{G}}}{\mu_{\mathcal{G}\mathcal{G}}}\right) + \widetilde{\kappa}_{1\mathcal{G}\mathcal{G}_{pdf}} \right) \\ &- \frac{L_{\mathcal{G}\mathcal{G}}}{\mu_{\mathcal{G}\mathcal{G}}^{\eta_{\mathcal{G}\mathcal{G}}}} \int_0^{\infty} \mathcal{G}\mathcal{G}_{pdf}(x) x^{\eta_{\mathcal{G}\mathcal{G}}} dx \end{aligned}$$

where the term  $\int_0^{\infty} \mathcal{G}\mathcal{G}_{pdf}(x) x^{\eta_{\mathcal{G}\mathcal{G}}} dx$  corresponds to the classical moment of order  $\eta_{\mathcal{G}\mathcal{G}}$  of the Generalized Gamma law and can be expressed as [10]:

$$\begin{aligned} &\int_0^{\infty} \mathcal{G}\mathcal{G}_{pdf}(x) x^{\eta_{\mathcal{G}\mathcal{G}}} dx \\ &= \mu_{\mathcal{G}\mathcal{G}}^{\eta_{\mathcal{G}\mathcal{G}}} \end{aligned}$$

#### ACKNOWLEDGMENT

The authors would like to thank Dr Michelle M. Horta and Pr Nelson D. A. Mascarenhas (Federal University of São Carlos, Computer Department, São Carlos, Brazil) for their fruitful discussions and contributions to these works.

They would like to thank the reviewers for their suggestions to improve the quality of the paper and generalize the original contribution.

This work has been partially funded by ANR (the French National Research Agency) and DGA (Direction Générale de l'Armement) under ALYS project ANR-15-ASTR-0002.

#### REFERENCES

- [1] C. A. Deledalle, L. Denis, and F. Tupin, "Iterative Weighted Maximum Likelihood Denoising with Probabilistic Patch-Based Weights," *IEEE Transactions on Image Processing*, vol. 18, no. 12, pp. 2661–2672, 2009.
- [2] F. Galland, J.-M. Nicolas, H. Sportouche, M. Roche, F. Tupin, and P. Réfrégier, "Unsupervised synthetic aperture radar image segmentation using Fisher distributions," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 8, pp. 2966–2972, 2009.

- [3] C. Tison, J.-M. Nicolas, F. Tupin, and H. Maître, "A new statistical model for markovian classification of urban areas in high-resolution SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 10, pp. 2046–2057, 2004.
- [4] H. Sportouche, F. Tupin, J.-M. Nicolas, T. Perciano, and C. A. Deledalle, "How to combine TerraSAR-X and Cosmo-SkyMed high-resolution images for a better scene understanding?" in *International Geoscience and Remote Sensing Symposium (IGARSS 2012)*. IEEE, 2012, pp. 178–181.
- [5] H. Sportouche, F. Tupin, and L. Denise, "Extraction and three-dimensional reconstruction of isolated buildings in urban scenes from high-resolution optical and SAR spaceborne images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 10, pp. 3932–3946, 2011.
- [6] M. M. Horta, N. D. A. Mascarenhas, H. Sportouche, N. Seichepine, F. Tupin, and J.-M. Nicolas, "Change detection in multitemporal HR SAR images: A hypothesis test-based approach," in *International Geoscience and Remote Sensing Symposium (IGARSS 2012)*. IEEE, 2012, pp. 374–377.
- [7] F. Bujor, E. Trouvé, L. Valet, J.-M. Nicolas, and J.-P. Rudant, "Application of log-cumulants to the detection of spatiotemporal discontinuities in multitemporal SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 10, pp. 2073–2084, 2004.
- [8] J. W. Goodman, "Some fundamental properties of speckle," *Journal of the Optical Society of America*, vol. 66, no. 11, pp. 1145–1150, 1976.
- [9] J.-M. Nicolas, "Introduction aux statistiques de deuxième espèce : applications des logs-moments et des logs-cumulants à l'analyse des lois d'images radar," *Traitement du signal*, vol. 19, no. 3, pp. 139–167, 2002.
- [10] —, "Application de la transformée de Mellin : étude des lois statistiques de l'imagerie cohérente," Institut Telecom, Telecom ParisTech, TII Group, TSI Department, Paris, France, Tech. Rep. 2006D010, 2006.
- [11] Y. Delignon, R. Garello, and A. Hillion, "Statistical modelling of ocean SAR images," *IEE Proceedings Radar, Sonar and Navigation*, vol. 144, no. 6, pp. 348–354, 1997.
- [12] A. C. Frery, H.-J. Muller, C. d. C. F. Yanasse, and S. J. S. Sant'Anna, "A model for extremely heterogeneous clutter," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35, no. 3, pp. 648–659, 1997.
- [13] M. S. Majd, E. Simonetto, and L. Polidori, "Maximum likelihood classification of high-resolution SAR images in urban area," in *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences (ISPRS 2011)*, vol. XXXVIII-4/W19, 2011.
- [14] —, "Maximum likelihood classification of single high resolution polarimetric SAR images in urban areas," *Photogrammetrie-Fernerkundung-Geoinformation*, vol. 2012, no. 4, pp. 395–407, 2012.
- [15] E. W. Stacy, "A Generalization of the Gamma distribution," *The Annals of Mathematical Statistics*, vol. 33, no. 3, pp. 1187–1192, 1962.
- [16] E. W. Stacy and G. A. Mihram, "Parameter estimation for a generalized Gamma distribution," *Technometrics*, vol. 7, no. 3, pp. 349–358, 1965.
- [17] H.-C. Li, W. Hong, and Y.-R. Wu, "Generalized Gamma distribution with MoLC estimation for statistical modeling of SAR images," in *Proc. IEEE 1st Asian and Pacific Conference on Synthetic Aperture Radar (AP SAR 2007)*, 2007, pp. 525–528.
- [18] H.-C. Li, W. Hong, Y.-R. Wu, and P.-Z. Fan, "An efficient and flexible statistical model based on generalized Gamma distribution for amplitude SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 6, pp. 2711–2722, 2010.
- [19] —, "On the empirical-statistical modeling of SAR images with generalized Gamma distribution," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 3, pp. 386–397, 2011.
- [20] C. C. Freitas, A. C. Frery, and A. H. Correia, "The polarimetric G distribution for SAR data analysis," *Environmetrics*, vol. 16, no. 1, pp. 13–31, 2005.
- [21] S. Khan and R. Guida, "On single-look multivariate  $\mathcal{G}$  distribution for PolSAR data," *IEEE Journal of Selected Topics in Applied Earth and Remote Sensing*, vol. 5, no. 4, pp. 1149–1163, 2012.
- [22] H. Sportouche, F. Tupin, and J.-M. Nicolas, "Capacités d'imitation de la loi Gamma Généralisée pour la modélisation statistique d'images radar haute résolution," in *XXVe Colloque GRETSI (Groupe d'Études du Traitement du Signal et des Images) (GRETSI 2015)*, 2015.
- [23] —, "Mimick capacity of Generalized Gamma distribution for high resolution SAR image statistical modeling," in *International Geoscience and Remote Sensing Symposium (IGARSS 2016)*. IEEE, 2016.
- [24] I. R. Joughin, D. B. Percival, and D. P. Winebrenner, "Maximum likelihood estimation of K distribution parameters for SAR data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 31, no. 5, pp. 989–999, 1993.
- [25] Z. Sun, C. Han, and X. Yuan, "Parameter estimation of K distribution based on second-kind statistics," in *Proc. IEEE 13th Conference on Information Fusion (FUSION 2010)*, 2010, pp. 1–6.
- [26] S. N. Anfinsen and T. Eltoft, "Application of the matrix-variate Mellin transform to analysis of polarimetric radar images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 6, pp. 2281–2295, 2011.
- [27] S. N. Anfinsen, A. P. Doulgeris, and T. Eltoft, "Goodness-of-fit tests for multilook polarimetric radar data based on the Mellin transform," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 7, pp. 2764–2781, 2011.
- [28] V. A. Krylov, G. Moser, S. B. Serpico, and J. Zerubia, "On the method of logarithmic cumulants for parametric probability density function estimation," INRIA, Vision, Perception and Multimedia Understanding, Sophia Antipolis, France, Tech. Rep. 7666, 2011.
- [29] —, "On the method of logarithmic cumulants for parametric probability density function estimation," *IEEE Transactions on Image Processing*, vol. 22, no. 10, pp. 3791–3806, 2013.
- [30] A. H. Zemanian, *Distribution Theory and Transform Analysis: An Introduction to Generalized Functions, with Applications*. Dover, 1965.
- [31] L. Bombrun, S. N. Anfinsen, and O. Harant, "A complete coverage of log-cumulant space in terms of distributions for polarimetric SAR data," in *5th International Workshop on Science and Applications of SAR Polarimetry and Polarimetric Interferometry (POLInSAR 2011)*, 2011.
- [32] J.-M. Nicolas, "A Fisher-MAP filter for SAR image processing," in *International Geoscience and Remote Sensing Symposium (IGARSS 2003)*, vol. 3. IEEE, 2003, pp. 1996–1998.
- [33] S. Kullback and R. A. Leibler, "On information and sufficiency," *The Annals of Mathematical Statistics*, vol. 22, no. 1, pp. 79–86, 1951.
- [34] I. N. Sanov, "On the probability of large deviations of random magnitudes," *Matematicheskii Sbornik*, vol. 84, no. 1, pp. 11–44, 1957.
- [35] T. M. Cover and J. A. Thomas, *Elements of information theory*. New York: Wiley-Interscience, 1991.
- [36] F. Johansson and others., "mpmath: a Python library for arbitrary-precision floating-point arithmetic (version 0.18)," <http://mpmath.org>, 2013.
- [37] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series and products*. Academic Press, 1980.