PEOD : Padé Estimated Optimum (radar) Detector

Emmanuelle Jay [1,2], Jean-Philippe Ovarlez [2], David Declercq [1], Patrick Duvaut [1,3]

[1] ENSEA/UCP-ETIS, URA D2235 - 6 av. du Ponceau, BP 44, 95014 Cergy Pontoise Cedex, France
[2] ONERA/DEMR/TSI - BP72, F92322 Châtillon Cedex, France
[3] Globespan Inc. Tech. - 100 Schulz Drive - Red Bank - NJ 07701 - USA
Email : jay@onera.fr, ovarlez@onera.fr, declercq@ensea.fr, pdu@globespan.net

Abstract— **In this paper, an expression of the optimum non-Gaussian radar detector is derived from the non-Gaussian SIRP model (Spherically Invariant Random Process) clutter and a Padé approximation of the** *characteristic function* **of the SIRP. The SIRP model is used to perform coherent detection and to modelize the non-Gaussian clutter as a complex Gaussian process whose variance is itself a positive random variable (r.v.). The probability density function (PDF) of the variance characterizes the statistics of the SIRP and after performing a Padé approximation of this PDF from reference clutter cells we derive the so-called Padé Estimated Optimum (Radar) Detector (PEOD) without any knowledge about the statistics of the clutter. We evaluate PEOD performance for an unknown target signal embedded in K-distributed clutter and compare with optimum detectors performance (optimum in particular clutter statistics such as Optimum K Detector - OKD - in K-distributed clutter).**

I. Introduction

Coherent radar detection against non-Gaussian clutter has gained many interests in the radar community since experimental clutter measurements made by organizations like MIT [1] have shown to fit non-Gaussian statistical models. One of the most tractable and elegant non-Gaussian model results in the so-called *Spherically Invariant Random Process* (SIRP) theory which states that non-Gaussian random vector is the product between Gaussian random vector with a non-negative random variable (r.v.) (the variance of the Gaussian process is itself a r.v.). This model allows to derive non-Gaussian joint probability density function (PDF) and then optimum radar detector strategies. For example in [2], the optimum radar detector is derived in the presence of composite disturbance of known statistics modeled as SIRP.

The goal of this work consists in estimating the variance PDF of the noise with a Padé approximation. This allows to derive in a closed form the joint PDF of the SIRV and to perform, from the likelihood ratio test (LRT), the Padé Estimated Optimum Detector (PEOD) for non-fluctuating and unknown target signal. Padé approximation is performed after *a posteriori* variance resampling which means to deal directly with the received clutter data. It is no more necessary to have any knowledge about the clutter statistics.

II. General relations of detection theory

A. Likelihood ratio test

We consider here the basic problem of detecting the presence (H_1) or absence (H_0) of a complex signal **s** in a set of N measurements of m-complex vectors $\mathbf{y} = \mathbf{y}_\mathrm{I} + \mathrm{j}\,\mathbf{y}_\mathrm{Q}$ corrupted by a sum **c** of independent additive complex noises (noises + clutter). The problem can be described in terms of a statistical hypothesis test :

$$H_0 : \mathbf{y} = \mathbf{c} \quad (1)$$

$$H_1 : \mathbf{y} = \mathbf{s} + \mathbf{c} \quad (2)$$

When present, the target signal **s** corresponds to a modified version of the perfectly known transmitted signal **t** and can be rewritten as $\mathbf{s} = A\,T(\underline{\theta})\,\mathbf{t}$. A is the target amplitude and we suppose determined all the other parameters $\underline{\theta}$ which characterize the target after the transformation T (Doppler frequency, time delay, ...). In the following, $\mathbf{p} = T(\underline{\theta})\,\mathbf{t}$. The observed vector **y** is used to form the LRT $\Lambda(\mathbf{y})$ which is compared to a threshold η in order to reach a desired false alarm probability (P_{fa}) value :

$$\Lambda(\mathbf{y}) = \frac{p_\mathbf{y}(\mathbf{y}/H_1)}{p_\mathbf{y}(\mathbf{y}/H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \quad (3)$$

LRT performance follow from the statistics of the data.
P_{fa} is the probability of choosing H_1 when the target is absent, and the detection probability (P_d) is the probability of choosing H_1 when the target is present, that is :

$$P_{fa} = \mathbb{P}(\Lambda(\mathbf{y}) \underset{H_0}{>} \eta) \quad \text{and} \quad P_d = \mathbb{P}(\Lambda(\mathbf{y}) \overset{H_1}{>} \eta), \quad (4)$$

where $\mathbb{P}(X > a)$ is the probability of X being greater than a under the PDF of the r.v. X.

B. Gaussian clutter case

When the clutter **c** is supposed to be complex Gaussian-distributed ($\mathcal{CN}(\mathbf{0}, 2\sigma^2\mathbf{M})$), the so-called Optimum Gaussian Detector (OGD) gives from (3) and for a known signal **s** :

$$\Re(\mathbf{y}^\dagger\,\mathbf{M}^{-1}\mathbf{s}) \underset{H_0}{\overset{H_1}{\gtrless}} \sigma^2\,\lambda + \frac{\mathbf{s}^\dagger\,\mathbf{M}^{-1}\mathbf{s}}{2}. \quad (5)$$

$\lambda = \log(\eta)$, $\Re(z)$ denotes the real part of z and † is the transpose conjugate operator.

When the target signal **s** is unknown and non-fluctuating, the generalized likelihood ratio test (GLRT) yields to :

$$|\mathbf{p}^\dagger \, \mathbf{M}^{-1} \, \mathbf{y}|^2 \underset{H_0}{\overset{H_1}{\gtrless}} 2\sigma^2 \lambda \, \mathbf{p}^\dagger \mathbf{M}^{-1} \mathbf{p}. \tag{6}$$

Statistic tests under H_0 and H_1 are respectively exponential and Rice-Nagakami-distributed and corresponding P_{fa} and P_d expressions are given in [3]. For fluctuating target, P_d expression has to be integrated over the fluctuation PDF, generally Swerling-K fluctuations (K is an integer), and expressions are given in [3].

C. Non-Gaussian clutter case - SIRV model

In the case of non-Gaussian clutter, detection strategies can be derived if we consider a particular clutter nature, i.e. if an *a priori* hypothesis is made on the clutter statistic.
Non-Gaussian clutter and general radar detector expressions come from the SIRP representation ([4], [5], [6]).
SIRV (Vector) model interpretes each element of the clutter vector **c** as the product of a m-complex Gaussian vector **x** ($\mathcal{CN}(\mathbf{0}, \in \mathbf{M})$) with a positive r.v. τ, that is $\mathbf{c} = \mathbf{x}\,\sqrt{\tau}$.

The PDF of the variable τ is called the *characteristic function* of the SIRV and the so formed vector **c** is, conditionnally to τ, a complex Gaussian random process ($\mathcal{CN}(\mathbf{0}, \in \tau\,\mathbf{M})$) with joint PDF $p(\mathbf{c}/\tau)$. The joint PDF of the SIRV gives :

$$p(\mathbf{c}) = \int_0^{+\infty} \frac{\tau^{-m}}{(2\pi)^m |\mathbf{M}|} \exp\left(-\frac{\mathbf{c}^\dagger \mathbf{M}^{-1} \mathbf{c}}{2\tau}\right) p(\tau) d\tau. \tag{7}$$

From (7), $p_{\mathbf{y}}(\mathbf{y}/H_0) = p_{\mathbf{c}}(\mathbf{y})$ and $p_{\mathbf{y}}(\mathbf{y}/H_1) = p_{\mathbf{y}}(\mathbf{y} - \mathbf{s}/H_0)$ for a known target signal **s**. The LRT becomes in this case :

$$\int_0^{+\infty} \left[\exp\left(-\frac{q_1(\mathbf{y})}{2\tau}\right) - \exp\left(\lambda - \frac{q_0(\mathbf{y})}{2\tau}\right)\right] \frac{p(\tau)}{\tau^m} \, d\tau \underset{H_0}{\overset{H_1}{\gtrless}} 0, \tag{8}$$

where $q_0(\mathbf{y}) = \mathbf{y}^\dagger \mathbf{M}^{-1} \mathbf{y}$, $q_1(\mathbf{y}) = q_0(\mathbf{y} - \mathbf{s})$ for a known signal **s** and $\lambda = \ln(\eta)$.

When the target signal **s** is unknown, the detection strategy is given by (8) where now :

$$q_1(\mathbf{y}) = \mathbf{y}^\dagger \mathbf{M}^{-1} \mathbf{y} - \frac{|\mathbf{p}^\dagger \mathbf{M}^{-1} \mathbf{y}|^2}{\mathbf{p}^\dagger \mathbf{M}^{-1} \mathbf{p}}. \tag{9}$$

D. Optimum K Detector (OKD)

In the case of K-distributed clutter (size m) with parameters ν and b, the r.v. τ is Gamma-distributed with parameters ν and $\beta = 2/b^2$. Following the same processes with (8), the expression of the so-called Optimum K-distributed Detector (OKD) becomes $\forall m \geq 2$:

$$\left(\frac{q_1(\mathbf{y})}{q_0(\mathbf{y})}\right)^{\frac{\nu - m}{2}} \cdot \frac{K_{\nu - m}\left(b\,\sqrt{q_1(\mathbf{y})}\right)}{K_{\nu - m}\left(b\,\sqrt{q_0(\mathbf{y})}\right)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta, \tag{10}$$

where $q_0(\mathbf{y}) = \mathbf{y}^\dagger \, \mathbf{M}^{-1} \mathbf{y}$ and $q_1(\mathbf{y})$ is given by (9) for unknown signal **s**.
For $m = 1$ the expression is given in [3].

Padé PDF approximation of $p(\tau)$ provides an expression in terms of a sum of weighted complex decaying exponential which allows to perform integration over $p(\tau)$ in (8) to give the so-called Padé Estimated Optimum (radar) Detector (PEOD). Padé approximation method is presented in details in [7], [8]. In the next section we just recall the resulting expressions of PDF and CDF (cumulative density function) approximation.

III. PADÉ ESTIMATED OPTIMUM (RADAR) DETECTOR (PEOD)

Padé approximation method provides the following PDF and cumulative density function (CDF) expressions :

$$\tilde{p}(z) = \sum_{k=1}^{M} \lambda_k \, e^{-\alpha_k \, z} \quad \text{and} \quad \tilde{F}(z) = 1 - \sum_{k=1}^{M} \frac{\lambda_k}{\alpha_k} \, e^{-\alpha_k \, z}. \tag{11}$$

Padé coefficients $\{\alpha_k\}_{k=1}^M$ and $\{\lambda_k\}_{k=1}^M$ are in pairs conjugate when complex or real and all $\{\alpha_k\}$ real parts are positive (to assure the PDF convergence towards zero when z tends to $+\infty$).

When $p(\tau)$ is replaced by $\tilde{p}(\tau)$ in (8), integration becomes henceforth tractable [9] and yields to the PEOD expression :

$$\left(\frac{q_1(\mathbf{y})}{q_0(\mathbf{y})}\right)^{\frac{1-m}{2}} \cdot \frac{\sum_{k=1}^{M} \lambda_k (\alpha_k)^{\frac{m-1}{2}} K_{1-m}\left(\sqrt{B_k^1(\mathbf{y})}\right)}{\sum_{k=1}^{M} \lambda_k (\alpha_k)^{\frac{m-1}{2}} K_{1-m}\left(\sqrt{B_k^0(\mathbf{y})}\right)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta, \tag{12}$$

where $q_0(\mathbf{y}) = \mathbf{y}^\dagger \, \mathbf{M}^{-1} \, \mathbf{y}$, $q_1(\mathbf{y}) = q_0(\mathbf{y} - \mathbf{s})$ for **s** known, $q_1(\mathbf{y})$ is given by (9) for **s** unknown, $B_k^j(\mathbf{y}) = 2\,\alpha_k q_j(\mathbf{y})$, $j = 0, 1$.

Based on moment generating function (MGF) approximation, Padé approximation method requires knowledge of the moments of the r.v. τ up to the order $L + M + 1$ or estimation of the moments from samples of $p(\tau)$ (the integer L is less than M (generally $L = M - 1$) and needed to perform Padé approximation. See [7], [8] for more details). But neither true moments nor samples from $p(\tau)$ are available.

In the next section we propose to regenerate variance samples according to the *a posteriori* PDF (APDF) of the variance (from reference clutter cells and with a non-informative variance prior) and to perform Padé approximation with the estimated moments of the resamples.

We have made comparisons between the two approaches via simulations and it can be shown that errors due to the moments estimation does not involve significant loss of performance in PEOD.

A. *Variance resampling*

From N_r reference clutter cells of size m, $\mathbf{R} = [\mathbf{r}_1, \cdots, \mathbf{r}_{N_r}]^T$ where $\mathbf{r}_i = [r_i(1), \cdots, r_i(m)]^T$ and with the Bayes'rule, we regenerate N_r variance samples according to the variance APDF.
The Bayes'rule provides us directly the APDF :

$$p(\tau/\mathbf{r}_i) = \frac{p(\mathbf{r}_i/\tau)g(\tau)}{p(\mathbf{r}_i)}. \tag{13}$$

$p(\mathbf{r}_i)$ is the normalization constant calculated in integrating the numerator of (13) over τ :

$$p(\mathbf{r}_i) = \int_0^{+\infty} p(\mathbf{r}/\tau_i)g(\tau)d\tau. \tag{14}$$

$g(\tau)$ is the prior density of the variance for the reference clutter cells. As we do not have any knowledge (except the variance positivity) about the variance density, we choose a non-informative prior, called Jeffreys' prior, which is proportional to the square root of Fisher's information measure.
So, we have :

$$g(\tau) = \frac{1}{\tau}. \tag{15}$$

After calculation, the expression of the variance APDF becomes an inverse gamma PDF ($\mathcal{IG}$ PDF) and the N_r variance are resampled according to :

$$\tau_{i=1}^{N_r} \sim \mathcal{IG}\left(m, \frac{2}{\mathbf{r}_i^\dagger \mathbf{M}^{-1}\mathbf{r}_i}\right) \tag{16}$$

where $\mathcal{IG}(.)$ is the inverse Gamma distribution (i.e. the PDF of the inverse of a gamma r.v.).

We can also choose a *conjugate* prior density in our case because of the form of the likelihood of the data.
A prior density is called *conjugate* if the resulting APDF own to the same PDF family than the prior density. So, we could choose an inverse gamma PDF for the prior density with parameters a_p and b_p (the subscript p is for prior) :

$$g(\tau) = \frac{1}{b_p^{a_p}\Gamma(a_p)}\tau^{-a_p-1}e^{-\frac{1}{\tau b_p}}, \tag{17}$$

and the N_r variance are so resampled according to :

$$\tau_{i=1}^{N_r} \sim \mathcal{IG}\left(m + a_p, \frac{2b_p}{2 + b_p\mathbf{r}_i^\dagger \mathbf{M}^{-1}\mathbf{r}_i}\right). \tag{18}$$

We have to choose a_p and b_p values to keep the prior density as non-restrictive as possible.
On figures (1) and (2) different $\mathcal{IG}$ PDF are plotted for different values of the parameters a_p and b_p.
If $a_p \longrightarrow 0$ and $b_p \longrightarrow +\infty$ then $\mathcal{IG}(\tau; a_p, b_p) \propto a_p/\tau$. That is, the conjugate prior density tends to a non-informative density.
This is shown on figure (3) and we choose $a_p = 10^{-3}$ and $b_p = 3.65$.

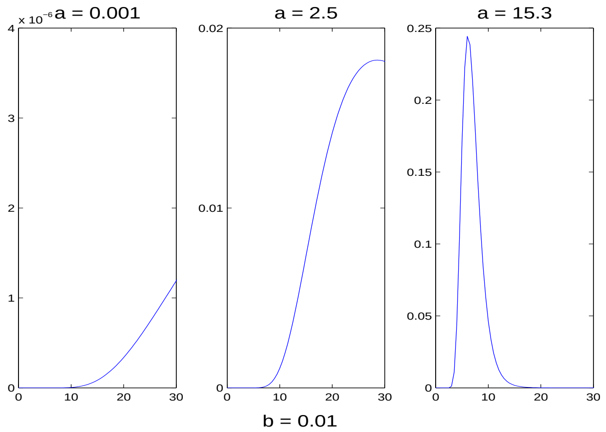

Fig. 1. Inverse gamma(x;a_p,b_p) PDF for $b_p = 0.01$ and $a_p = 0.001, 2.5, 15.3$.

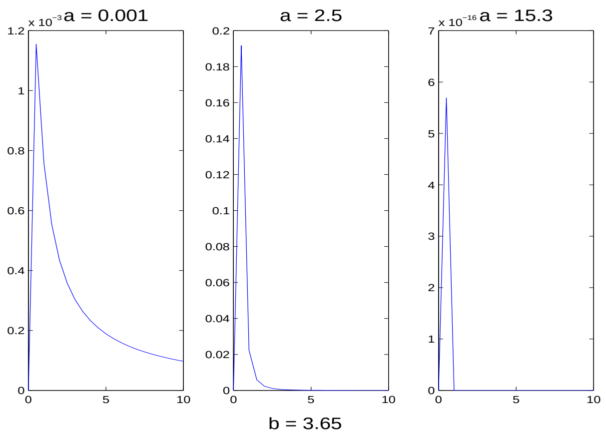

Fig. 2. Inverse gamma(x;a_p,b_p) PDF for $b_p = 3.65$ and $a_p = 0.001, 2.5, 15.3$.

All the $L + M + 1$ moments are then computed empirically from which M Padé coefficients $\{\alpha\}_k$ and $\{\lambda\}_k$ are derived in order to perform PEOD.
Variance resampling allows to deal directly with the data without statistical assumption on the clutter and PEOD performances are shown below. PEOD expression stands even if the

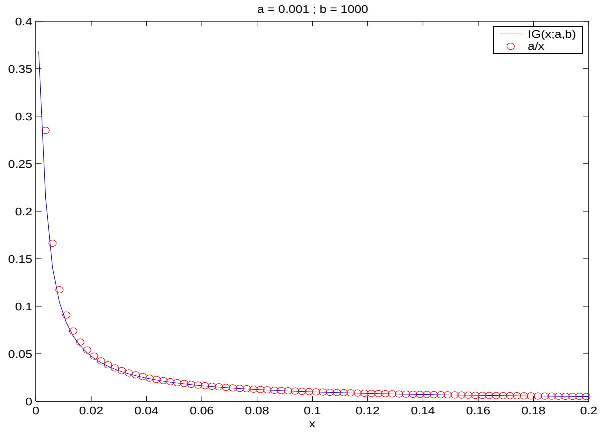

Fig. 3. Comparison between IG(x;a_p,b_p) and a/x for small a_p value ($a_p = 10^{-3}$) and high b_p value ($b_p = 10^3$).

characteristic function $p(\tau)$ of the SIRV is theoritically unknown like for Weibull clutter.
This is shown on **figure (6)**.

IV. Simulations

We compare the performances of the PEOD with those of OGD and OKD.
K-distributed clutter is generated with different values of the form parameter $\nu = 0.5,\ 20$. Smaller is the value of ν and spikier is the clutter. Inversely, when ν is high, K-PDF tends to a Gaussian distribution and this fact is confirmed through OGD and OKD performances on **figure (5)**.
Moreover, from PEOD expression, we can deduce OGD and OKD expression for particular variance PDF.
On **figure (6)**), we compare PEOD, OGD and OKD performances (OKD with different values of ν) for an unknown target signal embedded in Weibull clutter.

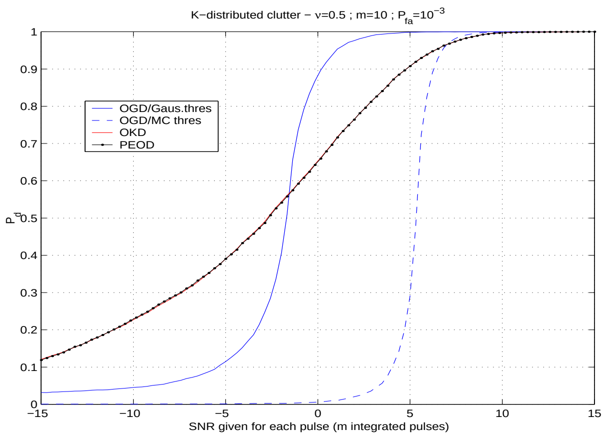

Fig. 4. Performances comparison between the OGD, OKD and PEOD for K-distributed clutter ($\nu = 0.5$, $P_{fa} = 10^{-3}$, $m = 10$ integrated pulses)

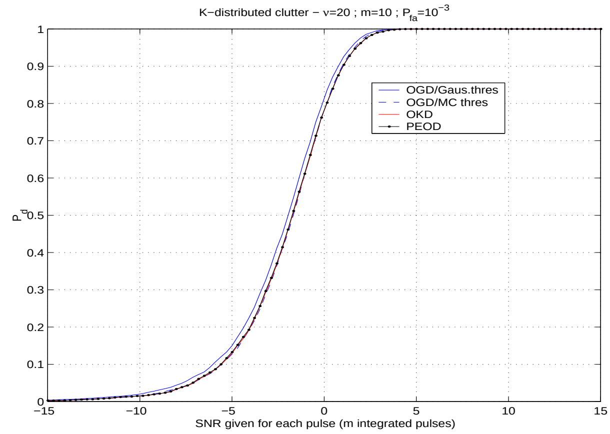

Fig. 5. Performances comparison between the OGD, OKD and PEOD for K-distributed clutter ($\nu = 20$, $P_{fa} = 10^{-3}$, $m = 10$ integrated pulses)

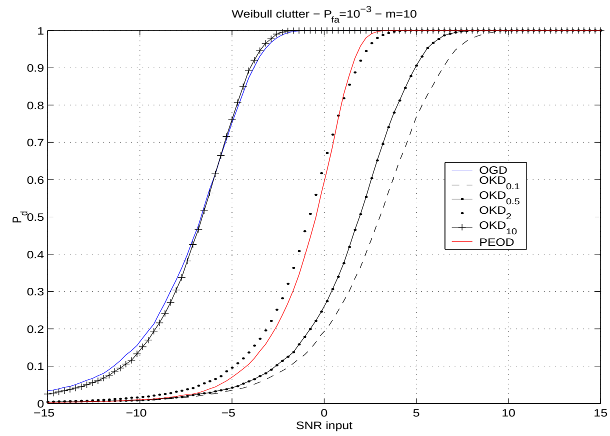

Fig. 6. Performances comparison between the OGD, OKD and PEOD for an unknown target signal in Weibull-distributed clutter ($a = 0.2$, $b = 2$; $P_{fa} = 10^{-3}$, $m = 10$ integrated pulses and $\nu = 0.1, 0.5, 2, 10$ for OKD). SNR given on x-axis is for one pulse : 0 dB before coherent post-integration corresponds to $10 \log 10(m)$ dB after the coherent pulse integration.

V. Conclusions

The present paper has addressed the contribution of Padé approximation method to the problem of coherent radar detection of a target embedded in a clutter with unknown statistics. The simple expression of PEOD (Padé Estimated Optimum Detector) allows to build the optimum radar detector and to evaluate its detection performances without having the knowledge of the clutter statistic.

VI. Outlooks : improving PEOD

At that time we are now able to give an improvement to PEOD, that is to say, to avoid the Padé approximation step of the variance PDF in estimating this PDF thanks to a bayesian estimator from reference clutter cells and the same variance prior density.

The resulting expression is called BORD (for Bayesian Optimum Radar Detector) and will be found in further papers.

BORD is "self-adaptative" to the clutter statistics because of its only dependence of the data (references and observed data). BORD performances reach OKD performances for K-distributed clutter and can apply for any clutter statistics (see on figures (7) and (8)). The only hypothesis made at the beginning of the study is the clutter SIRP modelization.

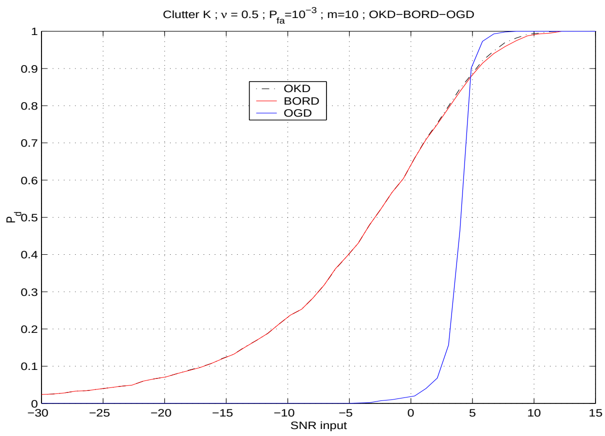

Fig. 7. Performances comparison between the OGD, OKD and BORD for K-distributed clutter ($\nu = 0.5$, $P_{fa} = 10^{-3}$, $m = 10$ integrated pulses).

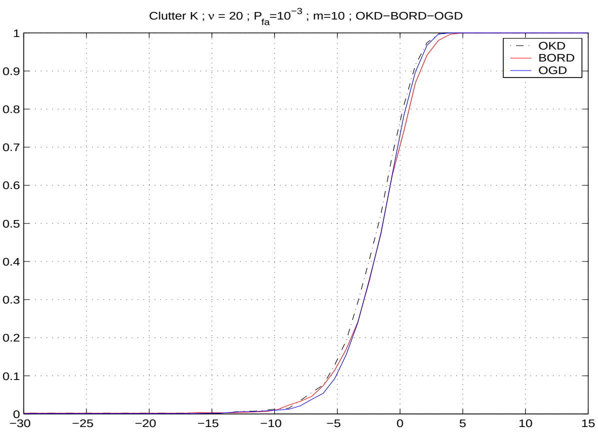

Fig. 8. Performances comparison between the OGD, OKD and BORD for K-distributed clutter ($\nu = 20$, $P_{fa} = 10^{-3}$, $m = 10$ integrated pulses).

References

[1] J.B. Billingsley, *Ground Clutter Measurements for Surface-Sited Radar*, *Technical Report 780, MIT*, February 1993

[2] F. Gini, M. V. Greco ,A. Farina and P. Lombardo, *Optimum and mismatched detection against K-distributed clutter plus Gaussian clutter*, *IEEE Trans.-AES* , **34**–3 (July 1998) 860–876.

[3] E. Jay, J.P. Ovarlez , D. Declercq and P. Duvaut, *Evaluation of Radar Detection Performances in Low Grazing Angle Clutter Environment*, *SET Panel Symposium on Low Grazing Angle Clutter* , Laurel (MD) USA, April 2000

[4] T. J. Barnard, and D. D. Weiner, *Non-Gaussian clutter modeling with generalized spherically invariant random vectors*, *IEEE Trans.-SP* , **44**–10 (October 1996) 2384–2390.

[5] M. Rangaswamy, D. Weiner and A. Öztürk, *Non-Gaussian Random Vector Identification Using Spherically Invariant Random Processes*, *IEEE Trans.-AES* **29**–1 (January 1993) 111–123

[6] K. Yao, *A representation theorem and its application to spherically invariant random processes*, *IEEE Trans.-IT* **19**–2 (1973) 600–608.

[7] H. Amindavar and J. A. Ritcey, *Padé Approximations of Probability Density Functions*, *IEEE Trans.-AES* **30**–2 (April 1994) 416–424.

[8] E. Jay, J.P. Ovarlez and P. Duvaut, *New Methods of Radar Performances Analysis*, *Signal Processing* **80**–12 (December 2000) 2527-2540.

[9] A. Erdélyi, Tables of integral transforms, *Bateman manuscript project, California Institute of Technology McGraw-Hill* , **1-2**– (1954).

[10] C. Robert, Méthodes de Monte Carlo par Chaînes de Markov, *Statistique mathématique et probabilité, Economica, Paris* (1996)

[11] M. A. Tanner, Tools for Statistical Inference, *Springer Series in Statistics, Springer-Verlag, N.-Y.* (1996)