A Least Informative Distribution of Ranging Errors in Robust Estimation of Localization

Georgy Shevlyakov
Peter the Great St. Petersburg Polytechnic University
St. Petersburg, Russia
Georgy.Shevlyakov@phmf.spbstu.ru

Kiseon Kim
Gwangju Institute of Science and Technology
Gwangju, South Korea
kskim@gist.ac.kr

***Abstract*—In the framework of the Huber's minimax variance approach to designing robust estimates of localization parameters, a generalization of the classical least informative distributions minimizing Fisher information for location is obtained in the wide class of ranging error distributions with a bounded quantile value. The considered variational problem set up naturally originates from the real-life problem of estimation of the unknown coordinates of an asset surrounded by the beacons with known positions.**

I. Introduction

The problems of accurate and reliable estimation of localization (location) are primary both in numerous applications (the localization of people and assets, the navigation beyond GPS coverage, etc.), especially in harsh environments (say, in the non-line-of-sight (NLOS) condition), as well as in the theory and applications of statistical signal processing [1], [2], [3]. The balance between accuracy and reliability, that is, between efficiency and stability of statistical procedures, is achieved by the methods of robust statistics [4], [5], [6], [7].

Recall that robustness of a statistical procedure means its stability to possible uncontrolled departures from the accepted probabilistic models of the underlying data distribution. This branch of mathematical statistics has been definitely formed as a robust extension of the classical parametric statistics by the end of the eighties, in particular, this mostly refers to robust estimation of location.

In present, there exist two general approaches to the design of robust estimates: historically the first Huber's minimax approach of quantitative robustness [4], [5], [6] and the Hampel's local approach based on influence functions (qualitative robustness) [7]. Within the first approach, the least informative (favorable) distribution minimizing Fisher information in a given class of data distributions is used with the subsequent application of the maximum likelihood parameter estimate for this least informative distribution. In this case, the minimax approach provides the guaranteed accuracy of estimation, namely, the asymptotic variance of the minimax parameter estimate is upper bounded. Within the second approach, a parameter estimate is defined by the desired influence function, which suits the set of the qualitative measures of robustness of estimates such as their sensitivity to the presence of gross outliers in the data, to the data rounding-off, etc.

In what follows, we mostly deal with the Huber's minimax approach to the design of robust parameter estimates. Recall that, generally, the minimax approach aims at the best solution in the worst situation providing a guaranteed result, sometimes rather pessimistic. However, Huber's implementation of the minimax principle in the problem of robust estimation luckily gives quite reasonable results.

The key-step of this approach to robust parameter estimation is the solution of the variational problem of minimizing Fisher information in a given class of pdfs (with a bounded distribution variance, or a bounded distribution quantile, or a bounded pdf, or a bounded distribution cumulative function, etc.)—next we briefly review several classical results on the least informative distributions, some of them work in further novel constructions.

An outline of the remainder of the paper is as follows. In Section II, the Huber's minimax approach is briefly reviewed. In Section III, a general problem set up for the design of minimax variance *M*-estimates of localization is given. In Section IV, the least informative distribution of the ranging errors is derived. In Section V, some conclusions are drawn.

II. Huber's minimax approach to robust estimation of location

A. Preliminaries

All further considered estimates belong to the class of M-estimates of location—now we define them. Let $x_1, \dots, x_n$ be a random sample from a distribution with pdf $f(x-\theta)$, where θ is a location parameter to be estimated. Next, assume that f belongs to a convex class $\mathcal{F}$ of symmetric and unimodal pdfs. Without any loss of generality, set $\theta = 0$.

An M-estimate T_n of the location parameter θ is defined as a solution to the following implicit equation

$$\sum_{i=1}^{n} \psi(x_i - T_n) = 0, \tag{1}$$

where $\psi(u)$ is an estimating or a score function.

Consider the following particular cases of (1):

1) for $\psi(u) = u$, we have the sample mean $\overline{x}_n$ as the estimate;
2) for $\psi(u) = sgn(u)$, we arrive at the sample median $med\, x$ as the estimate;
3) for a given pdf f, the choice $\psi(u) = -\log f(u)'$ $= -f'(u)/f(u)$ yields the maximum likelihood *(ML)* estimate.

Under rather general conditions of regularity imposed on the class Ψ of score functions ψ and on the class $\mathcal{F}$ of distribution densities f [4], [5], M-estimates are consistent and asymptotically normal with the asymptotic variance

$$var(n^{1/2}T_n) = V(\psi, f) = \frac{\int \psi^2(x)f(x)\,dx}{\left(\int \psi'(x)f(x)\,dx\right)^2}. \qquad (2)$$

The asymptotic variance of M-estimates satisfy the minimax or saddle-point property [4], [5]

$$V(\psi^*, f) \le V(\psi^*, p^*) \le V(\psi, p^*), \qquad (3)$$
$$V(\psi^*, p^*) = \sup_{f\in\mathcal{F}} \inf_{\psi\in\Psi} V(\psi, p),$$

where $f^*(x)$ is the least informative pdf f^* that minimizes Fisher information for location $I(f)$ over the class $\mathcal{F}$

$$f^* = \arg\min_{f\in\mathcal{F}} I(f), \qquad I(f) = \int \left[\frac{f'(x)}{f(x)}\right]^2 f(x)\,dx. \quad (4)$$

From (3) and (4) it follows that the optimal score function is given by the maximum likelihood method for the least informative pdf f^*

$$\psi^*(x) = -f^*(x)'/f^*(x).$$

The right-hand part of inequality (3) is nothing but the Cramér-Rao inequality

$$V(\psi, f^*) \ge V(-f^{*\prime}/f^*, f^*)$$
$$= 1\Big/\int (f^*(x)'/f^*(x))^2 f^*(x)\,dx = \frac{1}{I(f^*)},$$

whereas its left-hand part provides the guaranteed accuracy of minimax estimation, namely, the upper bound upon the asymptotic variance of the optimal robust M-estimate of location with the score function $\psi^*(x)$

$$V(\psi^*, f) \le \frac{1}{I(f^*)}$$

for any pdf $f(x)$ in the class $\mathcal{F}$.

Now we represent the well-known Huber minimax solution for the class of ε-contaminated Gaussian distributions (Tukey's gross-error model):

Consider

$$\mathcal{F}_\varepsilon = \{f\colon f(x) \ge (1-\varepsilon)\varphi(x),\ 0 \le \varepsilon < 1\},$$

where $\varphi(x)$ is the standard Gaussian.

In this case, the least informative density f^* has the central Gaussian part with the exponential tails

$$f^*(x) = \begin{cases} (1-\varepsilon)\varphi(x), & |x| \le k, \\ A\exp(-B|x|), & |x| > k, \end{cases}$$

The corresponding optimal score function is linear bounded

$$\psi^*(x) = \max\{-k, \min\{x, k\}\},$$

where the constants A, B and k are determined from the conditions of normalization and smoothness, namely, continuity and differentiability of the optimal density at the points $x = \pm k$.

B. The class of pdfs with bounded interquantile ranges

This class of pdfs is defined as follows:

$$\mathcal{F}_\beta = \left\{f\colon \int_{-l}^{l} f(x)\,dx = \beta\right\}. \qquad (5)$$

The parameters l and β ($0 < \beta \le 1$) are given; the latter characterizes the degree of closeness of $f(x)$ to a finite pdf. The restriction on this class means that the inequality $|X| \le l$ holds with probability β.

Now we define the class $\mathcal{F}_\beta$ in a slightly different way, namely as the class with a given interquantile distribution range $IQR_\beta = F^{-1}(1-\beta/2) - F^{-1}(\beta/2)$

$$\mathcal{F}_\beta = \{f\colon IQR_\beta(f) = 2l\}. \qquad (6)$$

C. The least informative pdfs and optimal score functions in the class with bounded interquantile ranges

Now we write out the least informative pdfs and the corresponding score functions for the class $\mathcal{F}_\beta$. In the class of pdfs with bounded interquantile ranges, the least informative pdf has two branches: the cosine one in the central part and the exponential at the tails [5], [8], [9] (see Fig. 1)

$$f_\beta^*(x) = \begin{cases} A_1\cos^2(B_1 x) & \text{for} \quad |x| \le l, \\ A_2\exp(-B_2|x|) & \text{for} \quad |x| > l, \end{cases} \qquad (7)$$

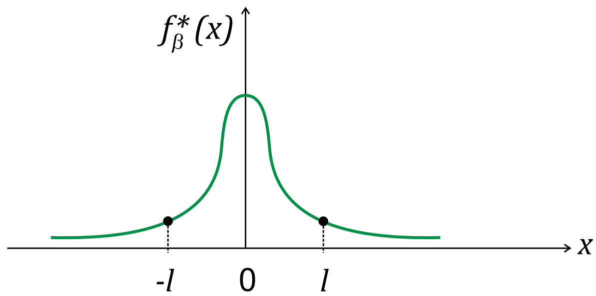

Fig. 1. The least informative pdf for the class $\mathcal{F}_\beta$

where the constants A_1, A_2, B_1, and B_2 are determined from the simultaneous equations characterizing the restrictions of the class $\mathcal{F}_\beta$, namely the conditions of normalization and boundness of interquantile ranges, and the conditions of continuity and continuous differentiability of the optimal solution at $x = l$ (for details, see [8], p. 64.)

The solution of those equations is given by the following formulas

$$A_1 = \frac{\beta\omega}{l(\omega + \sin(\omega))}, \qquad B_1 = \frac{\omega}{2l},$$
$$A_2 = \frac{(1-\omega)\lambda}{2l}e^\lambda, \qquad B_2 = \frac{\lambda}{l},$$

where the auxiliary parameters ω and λ satisfy the equations

$$\frac{2\cos^2(\omega/2)}{\omega\tan(\omega/2) + 2\sin^2(\omega/2)} = \frac{1-\beta}{\beta}, \qquad 0 < \omega < \pi$$

and $\lambda = \omega\tan(\omega/2)$.

The optimal score function $\psi^*_\beta(x)$ is bounded

$$\psi^*_\beta(x) = \begin{cases} \tan(B_1 x) & \text{for} \quad |x| \le l, \\ \tan(B_1 l) sgn(x) & \text{for} \quad |x| > l. \end{cases}$$

D. The least informative pdfs and optimal score functions in the classes of finite and non-degenerate distributions

Now we write out the least informative pdfs and the corresponding score functions for two particular cases of the class $\mathcal{F}_\beta$. First, consider the class of finite distributions $\mathcal{F}_1$:

$$\mathcal{F}_1 = \left\{ f \colon \int_{-l}^{l} f(x)\, dx = 1 \right\}. \tag{8}$$

This class corresponds to the case when the inequality $|X| \le l$ holds with probability 1. The additional restriction on this class yields the conditions of regularity at the boundaries $\pm l$

$$f(\pm l) = 0, \qquad f'(\pm l) = 0.$$

In this class, the least informative distribution density has the cosine form [8], [9] (see Fig. 2)

$$f^*_1(x) = \begin{cases} \dfrac{1}{l}\cos^2\left(\dfrac{\pi x}{2l}\right) & \text{for} \quad |x| \le l, \\ 0 & \text{for} \quad |x| > l. \end{cases} \tag{9}$$

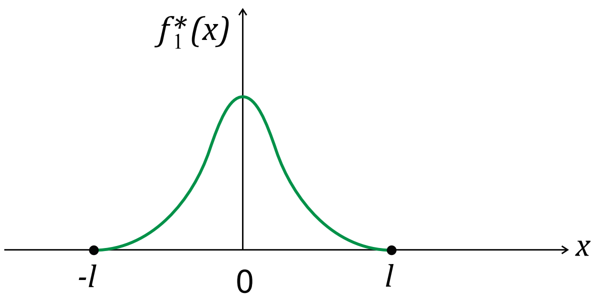

Fig. 2. The least informative pdf for the class $\mathcal{F}_1$

The optimal score function is unbounded: $\psi^*_1(x) = \tan\left(\dfrac{\pi x}{2l}\right)$ for $|x| \le l$.

Second, consider the class of non-degenerate distribution densities [8], [10]:

$$\mathcal{F}_0 = \left\{ f \colon f(0) \ge \frac{1}{2a} > 0 \right\}. \tag{10}$$

It is one of the most wide classes: any distribution density with a nonzero value at the center of symmetry belongs to it. The parameter a of this class characterizes the dispersion of the central part of a distribution.

In the class of nondegenerate distribution densities $\mathcal{F}_0$, the least informative distribution density is the Laplace or the double-exponential one [10] (see Fig. 3)

$$f^*_0(x) = L(x; 0, a) = \frac{1}{2a} \exp\left(-\frac{|x|}{a}\right) \tag{11}$$

with the optimal sign score function $\psi^*_0(x) = sgn(x)$.

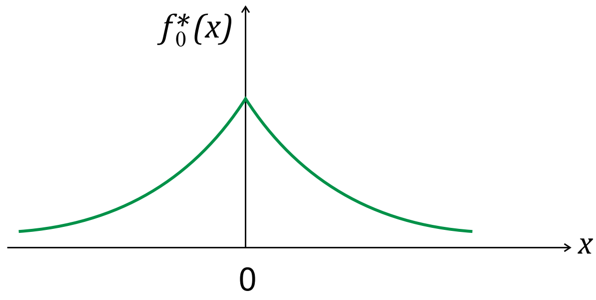

Fig. 3. The least informative pdf for the class $\mathcal{F}_0$

The least informative density in the class F_0 of nondegenerate distribution densities can be regarded as the limiting case of the optimal solution in the class $\mathcal{F}_\beta$ as the parameters $l \to 0$ and $\beta \to 0$ tend to zero so that

$$\frac{\beta}{2l} \to \frac{1}{2a}.$$

Now we comment on the aforementioned solutions. First, the exponential and trigonometric parts of the the least informative distribution densities arise due to the form of the solution of the basic variational problem of minimization of Fisher information for location $I(f) = \int (f'/f)^2 f\, dx$ under the side conditions of normalization $\int f\, dx = 1$ and nonnegativeness $f(x) = g^2(x) \ge 0$: the Euler-Lagrange equation for the new variable $g(x) = \sqrt{f(x)}$ is just the differential equation for the harmonic oscillator

$$4g''(x) + \lambda g(x) = 0,$$

where λ is the Lagrange multiplier corresponding to the side condition of normalization.

Second, the least informative density f^*_β also minimizes Fisher information in the class with the restriction of an inequality form

$$\int_{-l}^{l} f(x)\, dx \ge \beta, \qquad 0 < \beta \le 1$$

or as a class with an upper bound on the distribution interquantile range $IQR_\beta(f)$ of level β

$$IQR_\beta(f) \le 2l, \qquad 0 < \beta \le 1$$

as in the correctly posed problem the minimum of Fisher information is attained in the equality case.

III. Problem Statement and Motivation

A. Localization problem

Consider the localization problem setting given in [1]: there is a single agent with an unknown position $\mathbf{p} = (x, y, z)$ surrounded by N anchors with known positions $\mathbf{p}_i = (x_i, y_i, z_i)$, $i = 1, \dots, N$.

The distance between the agent and anchor i is denoted by $d_i(\mathbf{p}, \mathbf{p}_i) = ||\mathbf{p} - \mathbf{p}_i||_2 = \sqrt{(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2}$. The agent's estimates of this distance are $\hat{d}_i$, $i = 1, \dots, N$ with the ranging error $\Delta_i = \hat{d}_i - d_i(\mathbf{p}, \mathbf{p}_i)$.

Given the known anchor coordinates $\{(x_i, y_i, z_i)\}_1^N$ and the set of observations $\{\hat{d}_i\}_1^N$, we have to estimate the unknown localization coordinates (x, y, z).

B. Maximum likelihood estimates of localization

Generally, the pdf f_Δ of ranging errors is unknown, or partially known. First, we assume that f_Δ is given and the ranging errors $\{\Delta_i\}_1^N$ are i.i.d. random variables with pdf $f_\Delta(x) = f(x)$ defined on a non-negative support, that is, $\Delta_i \geq 0$. Then we can apply the maximum likelihood (*ML*) approach to estimating the unknown localization $\mathbf{p} = (x, y, z)$ given the observations $\{\mathbf{p}_i = (x_i, y_i, z_i)\}_1^N$ and $\{\hat{d}_i\}_1^N$.

The likelihood has the following form:

$$L(\hat{d}_1, \dots, \hat{d}_N; x, y, z) = \prod_1^N f(\Delta_i) = \prod_1^N f(\hat{d}_i - d_i(\mathbf{p}, \mathbf{p}_i));$$

$$\log L = \sum_1^N \log f(\hat{d}_i - d_i(\mathbf{p}, \mathbf{p}_i)), \tag{12}$$

which is to be maximized.

Next, we write out the estimating equations for localization $\mathbf{p} = (x, y, z)$, namely, the equations for the stationary point $(\hat{x}, \hat{y}, \hat{z})$ of the loglikelihood (12):

$$\frac{\partial \log L}{\partial \mathbf{p}} = \sum_1^N \frac{\partial \log f(\hat{d}_i - d_i(\mathbf{p}, \mathbf{p}_i))}{\partial \mathbf{p}}$$

$$= \sum_1^N \frac{\partial \log f(\Delta_i)}{\partial \Delta_i} \frac{(\hat{x} - x_i)\mathbf{i} + (\hat{y} - y_i)\mathbf{j} + (\hat{z} - z_i)\mathbf{k}}{d_i} = \mathbf{0},$$

which yields the following system of equations for $(\hat{x}, \hat{y}, \hat{z})$

$$\sum_1^N \frac{f'(\Delta_i)}{f(\Delta_i)} \frac{(\hat{x} - x_i)}{d_i} = 0,$$

$$\sum_1^N \frac{f'(\Delta_i)}{f(\Delta_i)} \frac{(\hat{y} - y_i)}{d_i} = 0,$$

$$\sum_1^N \frac{f'(\Delta_i)}{f(\Delta_i)} \frac{(\hat{z} - z_i)}{d_i} = 0$$

with the solutions in the form of the weighted averages

$$\hat{x} = \frac{\sum_1^N w_i x_i}{\sum_1^N w_i}, \; \hat{y} = \frac{\sum_1^N w_i y_i}{\sum_1^N w_i}, \; \hat{z} = \frac{\sum_1^N w_i z_i}{\sum_1^N w_i}, \tag{13}$$

where the weights w_i are

$$w_i = \frac{f'(\Delta_i)}{f(\Delta_i) d_i}, \quad i = 1, \dots, N.$$

For the solution of system (13), an iterative algorithm can be proposed when we begin with low-complexity initial estimates of localization coordinates $(\hat{x}^{(0)}, \hat{y}^{(0)}, \hat{z}^{(0)})$, say, with the sample means $(\bar{x}, \bar{y}, \bar{z})$ or the sample medians $(med\, x, med\, y, med\, z)$, and substitute them into the following iterative equations:

$$\hat{x}^{(k+1)} = \frac{\sum_1^N w_i(\hat{x}^{(k)}, \hat{y}^{(k)}, \hat{z}^{(k)}) x_i}{\sum_1^N w_i(\hat{x}^{(k)}, \hat{y}^{(k)}, \hat{z}^{(k)})},$$

$$\hat{y}^{(k+1)} = \frac{\sum_1^N w_i(\hat{x}^{(k)}, \hat{y}^{(k)}, \hat{z}^{(k)}) y_i}{\sum_1^N w_i(\hat{x}^{(k)}, \hat{y}^{(k)}, \hat{z}^{(k)})}, \tag{14}$$

$$\hat{z}^{(k+1)} = \frac{\sum_1^N w_i(\hat{x}^{(k)}, \hat{y}^{(k)}, \hat{z}^{(k)}) z_i}{\sum_1^N w_i(\hat{x}^{(k)}, \hat{y}^{(k)}, \hat{z}^{(k)})},$$

$$k = 0, 1, \dots, .$$

The convergence of algorithm (14) is usually provided: similar iterative schemes are widely used in robustness for computing *M*-estimates of location [5].

Under general conditions of regularity imposed on the ranging error pdfs [11], *ML*-estimates are consistent and asymptotically normal with the following covariance matrix

$$\mathbf{C} = \mathbf{I}^{-1},$$

where $\mathbf{I}$ is the Fisher information matrix of the diagonal form due to the i.i.d. assumption for ranging error random variables $\{\Delta_i\}_1^N$

$$\mathbf{I} = \begin{bmatrix} I_{11} & 0 & 0 \\ 0 & I_{22} & 0 \\ 0 & 0 & I_{33} \end{bmatrix}. \tag{15}$$

Now we compute the entries of matrix $\mathbf{I}$

$$I_{11}(x, y, z) = E\left(\frac{\partial \log L}{\partial x}\right)^2$$

$$= E \sum_1^N \left(\frac{f'(\Delta_i)}{f(\Delta_i)}\right)^2 \frac{(x - x_i)^2}{d_i^2}$$

$$= \sum_1^N E \left(\frac{f'(\Delta_i)}{f(\Delta_i)}\right)^2 \frac{(x - x_i)^2}{d_i^2}$$

$$= E \left(\frac{f'(\Delta_i)}{f(\Delta_i)}\right)^2 \sum_1^N \frac{(x - x_i)^2}{d_i^2}$$

$$= I(f) \sum_1^N \frac{(x - x_i)^2}{d_i^2},$$

where $I(f)$ is the Fisher information for location at the ranging error distribution

$$I(f) = \int_0^\infty \left(\frac{f'(x)}{f(x)}\right)^2 f(x)\, dx. \tag{16}$$

The similar structure arises for I_{22} and I_{33}:

$$I_{22} = I(f) \sum_1^N \frac{(y - y_i)^2}{d_i^2}, \quad I_{33} = I(f) \sum_1^N \frac{(z - z_i)^2}{d_i^2}.$$

Thus, the asymptotic accuracy of ML-estimates of the localization coordinates has the common factor $I(f)$. Then the covariance matrix of localization coordinate estimates is of the following form:

$$\mathbf{C} = \mathbf{I}^{-1}$$

$$= \frac{1}{I(f)} \begin{bmatrix} \sum \frac{(x - x_i)^2}{d_i^2} & 0 & 0 \\ 0 & \sum \frac{(y - y_i)^2}{d_i^2} & 0 \\ 0 & 0 & \sum \frac{(z - z_i)^2}{d_i^2} \end{bmatrix}^{-1}.$$

The structure of this obtained result allows for considering the minimax approach in robust estimation of location in some practically suitable class $\mathcal{F}$ of ranging error pdfs $f(x)$: $f \in \mathcal{F}$. Recall that the key-point of the minimax approach is the derivation of the least favorable (informative) pdf $f^*(x)$ minimizing Fisher information for location (16) over the chosen class with the subsequent use of the *ML*-estimate for this distribution density $f^*(x)$.

C. Problem setup

The choice of the class of ranging error pdfs strongly depends on the availability of experimental data on their distribution. Here we use the results of the extensive ranging measurement study performed in [1]. The obtained results on the histograms of ranging errors are different for the LOS and and NLOS conditions (see [1]). This information presented in Fig. 4 and Fig. 5 can be used in several possible ways:

1) designing an adaptive algorithm of estimation of localization by the use of the histogram (kernel density estimate) based *ML*-estimates;
2) designing an adaptive algorithm of estimation of localization by the use of a parametric pdf approximation of histograms with the subsequent application of *ML*-estimates;
3) designing a robust minimax algorithm of estimation of localization.

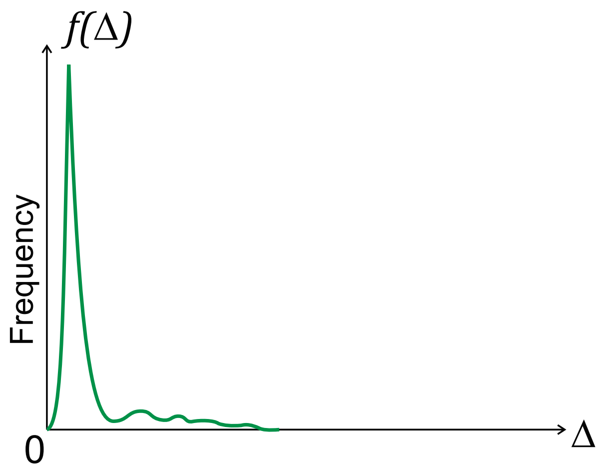

Fig. 4. Histogram of the ranging error for the LOS condition

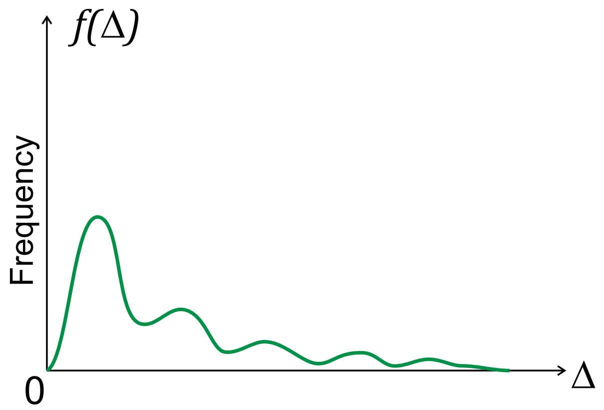

Fig. 5. Histogram of the ranging error for the NLOS condition

In this work, we follow way 3: in order to guarantee the quality of performance of estimation of localization, we propose to use Huber's minimax approach to robust estimation. The key-step in the implementation of the minimax approach to the problem of robust estimation is the solution of the variational problem of minimization of Fisher information over a certain class of pdfs. This class should be chosen basing on the experimentally observed specificity of the underlying distributions. Hence, we consider the class $\mathcal{F}_\Delta$ of pdfs with the non-negative support $f(x) \geq 0$ for $x \geq 0$ and $f(x) = 0$ otherwise, and with a bounded quantile value:

$$\int_0^l f(x)\,dx = \beta, \quad 0 < \beta < 1.$$

The parameter β is chosen, and the quantile value l is taken from the histogram

Finally, we pose the following variational problem:

$$f^* = \arg\min I(f) = \int_0^\infty \left[\frac{f'(x)}{f(x)}\right]^2 f(x)\,dx \qquad (17)$$

under the side conditions of non-negativity, normalization, bounded quantile value, and as it follows from Fig. 4 zero value of the pdf at the left boundary

1) $f(x) \geq 0, \quad \int_0^\infty f(x)\,dx = 1;$
2) $\int_0^\infty f(x)\,dx = \beta, \quad 0 < \beta < 1;$
3) $f(0) = 0.$

IV. MAIN RESULT

The solution of the aforementioned variational problem is given by the following statement.

Theorem The least informative (favorable) pdf for the class $\mathcal{F}_\Delta$ has the form (see Fig. 6):

$$f^*_\Delta(x) = \begin{cases} A_1 \cos^2(B_1(x - x_0)) & \text{for} \quad |x| \leq l, \\ A_2 \exp(-B_2 x) & \text{for} \quad |x| > l, \end{cases} \qquad (18)$$

where the constants A_1, A_2, B_1, B_2 and x_0 are determined from the simultaneous equations characterizing the restrictions of the class $\mathcal{F}_\Delta$.

These constants take the following values:

$$A_1 = \frac{2\beta\omega}{(\omega - \sin(\omega))l}, \qquad B_1 = \frac{\omega}{2l},$$

$$A_2 = \frac{(1-\beta)\lambda}{l} e^\lambda, \qquad B_2 = \frac{\lambda}{l},$$

$$x_0 = \frac{\pi l}{\omega},$$

where the auxiliary parameters λ and ω satisfy the equations

$$\lambda = -\frac{\omega}{\tan(\omega/2)},$$

$$\frac{\omega - \sin(\omega)}{2\omega \sin(\omega/2)^2} = \frac{\beta}{1-\beta},$$

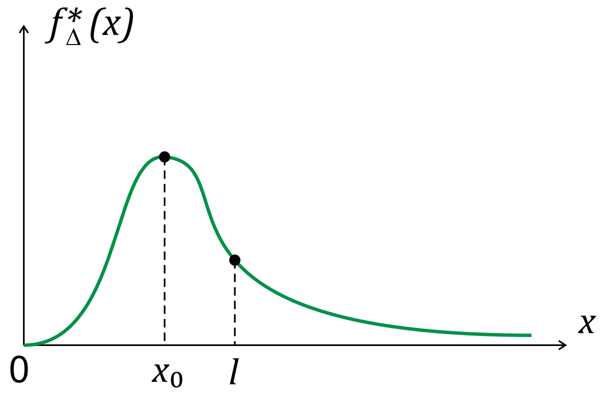

Fig. 6. The least informative pdf for the class $\mathcal{F}_\Delta$

where $\pi < \omega < 2\pi$.

The maximum likelihood score function $\psi^*_\Delta(x)$ is bounded (see Fig. 7)

$$\psi^*_\Delta(x) = \begin{cases} \tan(B_1 x) & \text{for} \quad |x| \le l, \\ \tan(B_1 l) sgn(x) & \text{for} \quad |x| > l. \end{cases}$$

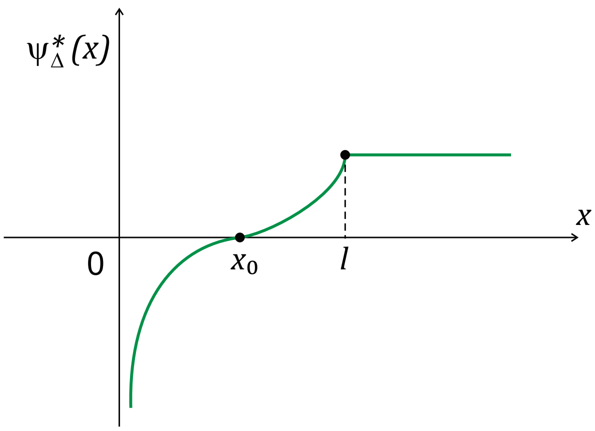

Fig. 7. The least informative pdf for the class $\mathcal{F}_\Delta$

The Fisher information at the least informative pdf takes the form:

$$I(f^*_\Delta) = \frac{\beta\omega^2}{l^2} + \frac{(1-\beta)\lambda^2}{l^2}.$$

Sketch of the proof. Basing on the similar minimization problem with symmetric distributions considered in Section II, we choose a candidate for the optimal solution in the form of the glued two principal extremals, trigonometric and exponential, of the basic variational problem. Here we have 5 unknown parameters: A_1, A_2, B_1, B_2 and x_0; the parameters l and β are given (taken from the histogram of the ranging errors in the line-of-site (LOS) and non-line-of-sight (NLOS) conditions, see Fig. 4 and Fig. 5).

The unknown parameters are determined from the conditions of normalization, given quantile value and smoothness:

$$\int_0^\infty f^*_\Delta(x)\,dx = 1, \qquad \int_0^l f^*_\Delta(x)\,dx = \beta,$$

$$f^*_\Delta(l-0) = f^*_\Delta(l+0), \quad (f^*_\Delta(l-0))' = (f^*_\Delta(l+0))',$$

$$f^*_\Delta(0) = 0.$$

It can be shown that the obtained result holds for the quantile levels lying in the interval $1/3 < \beta < 1$: this interval covers the central zone and the tail area of the ranging error distribution.

V. Conclusion

1) The novel result obtained in this work for the least informative pdf of ranging errors defined on the non-negative support generalizes the classical results obtained in the symmetric case. However, the existence of the least informative (favorable) distribution does not guarantee the existence of the minimax variance saddle-point for the problem of robust estimation of localization. So, this problem is open for the further study.
2) The finite sample size properties of the proposed algorithm also have to be studied, especially its bias.
3) A positive feature of the presented solution is that it holds for the most wide thinkable class of pdfs, namely, with a bounded quantile value: for instance, Pareto-type pdfs belong to this class.

Acknowledgment

The reported study was funded by RFBR according to the research project #18-29-03250.

References

[1] H. Wymeersch, S. Marano, W.M. Gifford, and M.Z. Win, “A machine learning approach to ranging error mitigation for UWB localization”, *IEEE Trans. Commun.*, vol. 60, Apr. 2012, pp. 1719-1728.

[2] H. Wymeersch, J. Lien, and M.Z. Win, “Cooperative localization in wireless networks”, *Proc. IEEE*, vol. 97, no. 2, Feb. 2009, pp. 427-450, special issue on *Ultra-Wide Bandwidth (UWB) Technology & Emerging Applications*.

[3] N. Patwari, J.N. Ash, S. Kyperountas, A.O. Hero, III, R.L. Moses, and N.S. Correal, “Locating the nodes: cooperative localization in wireless sensor networks”, *IEEE Signal Process. Mag.*, vol. 22, no. 4, Jul. 2005, pp. 54-69.

[4] P.J. Huber, ”Robust estimation of a location parameter”, *Ann. Math. Statist.*, vol. 35, no. 1, 1964, pp. 73-101.

[5] P.J. Huber, *Robust statistics*. New York: Wiley, 1981.

[6] P.J. Huber and E. Ronchetti, *Robust statistics*. New York: Wiley, 2009.

[7] F.R. Hampel, E. Ronchetti, P.J. Rousseeuw, and W.A. Stahel, *Robust Statistics. The Approach Based on Influence Functions*. New York: Wiley, 1986.

[8] G.L.Shevlyakov and N.O. Vilchevski, *Robustness in data analysis: criteria and methods*. Utrecht: VSP, 2002.

[9] G.L.Shevlyakov and N.O. Vilchevski, *Robustness in data analysis*. Boston: De Gruyter, 2011.

[10] B.T. Polyak and Ya.Z. Tsypkin, ”Robust identification”, *in Proc. 4th IFAC Symp.*, Oct. 1976, Tbilisi, pp. 203-224.

[11] H. Cramér, *Mathematical methods of statistics*. Princeton: Princeton University Press, 1974.