ABSTRACT
The Morosov’s discrepancy principle has been used as a general criterion to compute the regularization parameter in inverse problems. The Morosov’s principle is established when the discrepancy of the corresponding regularized solution is just equal to the measurement error. Considering the measurement error as a random variable, the goal of this paper is to present a generalized discrepancy principle for distributions in which the second moment is not defined. The generalized discrepancy principle is applied to several distributions: uniform, Gaussian, Cauchy, t-Student, Tsallis. The estimation of the initial condition in a heat transport problem is used as a test problem.

INTRODUCTION
The Morosov’s discrepancy principle [1-3] has been used as a general criterion to compute the regularization parameter in inverse problems. Essentially, the criterion is to calculate the root of the equation:

\[ K(x^{(\alpha(\delta))}) Y_0 = \delta \]  

being \( K(x) \) the forward model, \( Y_0 \) is the measured quantity, \( \delta \) is the error of the measurement, and \( \| \cdot \|_Y \) is a norm in the \( Y \) space. A heuristic procedure can be established to compute the root of equation (1) when the measured error is normally distributed (the probability density function is Gaussian), with zero mean and \( \sigma^2 \) variance (see [2], page 238). In this work a generalized discrepancy principle is presented, considering distributions in which the second moment is not defined. The generalized discrepancy principle is applied to several distributions: uniform, Gaussian, Cauchy, t-Student, Tsallis. The estimation of the initial condition in a heat transport problem is used as a test problem.

FORWARD TEST PROBLEM
The direct (forward) problem consists of a transient heat conduction problem in a slab with adiabatic boundary condition and initially at a temperature denoted by \( f(x) \). The mathematical formulation of this problem is given by the following heat equation

\[ \frac{\partial T(x, t)}{\partial t} = \frac{\partial^2 T(x, t)}{\partial x^2} \quad (x, t) \in \Lambda \times \mathbb{R}^+ \]

\[ \frac{\partial T(x, t)}{\partial x} = 0 \quad (x, t) \in \partial \Lambda \times \mathbb{R}^+ \]

\[ T(x, 0) = f(x) \quad (x, t) \in \Lambda \times \{0\} \]

where \( T(x, t) \) (temperature), \( f(x) \) (initial condition), \( x \) (spatial variable) and \( t \) (time variable) are dimensionless quantities and \( \Lambda = [0,1] \). The set of partial differential equations is solved by using a central finite difference approximation for space variable \( O(\Delta x^2) \), and explicit Euler method for numerical time integration \( O(\Delta t) \) [3].

The forward problem solution, for a given initial condition \( f(x) \), is explicitly obtained using separation of variables, for \((x, t) \in \Omega \times \mathbb{R}^+\):

\[ T(x, t) = \sum_{m=0}^{\infty} e^{-p_m^2 \Delta t} \frac{1}{N(\beta_m)} X(\beta_m, x) \int_0^t X(\beta_m, x') f(x') dx' \]

where \( X(\beta_m, x) = \cos(\beta_m x) \) are the eigenfunctions associated to the problem, \( \beta_m = m\pi \) are the eigenvalues, and \( N(\beta_m) = \int_0^1 X(\beta_m, x') f(x') dx' \).
represents the integral normalization (or the norm) \[4\]. The inverse problem consists in estimating the initial temperature profile \(f(x)\) for a given transient temperature distribution \(T(x,t)\) at time \(t\).

This problem has been extensively used for testing different methodologies in inverse problems [5–9], and it is badly conditioned problem [5].

**Inverse Analysis**

In general, inverse problems belong to the class of ill-posed problems, where existence, uniqueness and stability of their solutions cannot be ensured. Following the Tikhonov’s approach [10], a regularized solution is obtained by choosing the function \(f^*\) that minimizes the following functional

\[
J_a[f, f^*] = \| \tilde{T} - T(f) \|_2^2 + \alpha \Omega[f] \tag{3}
\]

where \(\tilde{T} = \tilde{T}(x,\tau)\) is the experimental data \(\tau = \tau\), \(T(f)\) is the temperature computed from the forward model at time \(\tau\), \(\Omega[f]\) denotes the regularization term given, \(\alpha\) is the regularization parameter, and \(\| \|_2\) is the 2-norm.

A scheme to determine the regularization parameter \(\alpha\) is the Morozov’s discrepancy principle: assuming that a bound \(\delta\) (or the ‘statistics’) of the measurement error is known, i.e., \(\| f_{\text{exact}} - \tilde{T} \|_2 \leq \delta\).

**The Morosov’s Discrepancy Principle**

For establishing a scheme to compute the regularization parameter, it is necessary to define the quantities: the residue \(R(f_0)\) and the error \(E(f_0)\) are defined by

\[
R(f_0) = \| \tilde{T} - T(f_0) \|_2^2 \tag{4}
\]

\[
E(f_0) = \| f - f_{\text{exact}} \|_2^2 \tag{5}
\]

The Morosov’s standard discrepancy principle is an a-posteriori parameter choice rule. It demands that a suitable regularized solution can be obtained under the condition:

\[
R(f_0) \cong N_0 \sigma^2 \tag{6}
\]

corresponding the optimum value for \(\alpha\) - the regularization parameter, and assuming that \(\sigma^2\) is the variance associate to a Gaussian distribution.

The last hypothesis can be justified by central limit theorem [11], and considering that the components of the random vector are uncorrelated (white Gaussian noise). The condition (6) is a particular case of the Morosov’s discrepancy principle.

**Optimization Algorithm**

The optimization problem is iteratively solved by the quasi-newtonian optimizer routine from the NAG Fortran Library [12], with variable metrics. This algorithm is designed to minimize an arbitrary smooth function subject to constraints (simple bound, linear or nonlinear constraints), using a sequential programming method.

This routine has been successfully used in several previous works: in geophysics, hydrologic optics, and meteorology.

**ESTIMATING INITIAL CONDITION**

Numerical experiments are carried out using two test functions, the triangular function

\[
f(x) = \begin{cases} 
2x & x \in [0, 0.5] \\
2(1-x) & x \in (0.5, 1] 
\end{cases} \tag{6}
\]

and semi-triangular function

\[
f(x) = \begin{cases} 
0.55 & 0 \leq x \leq 0.2 \\
8/3x & 0.2 < x < 0.5 \\
-28/5x + 23/5 & 0.5 < x < 0.75 \\
2/9 & 0 < x \leq 1 
\end{cases} \tag{7}
\]

In order to simulate the experimental data (measured temperatures at a time \(\tau > 0\)), which intrinsically should contain errors, a random perturbation is added to the exact solution of the direct problem, such that

\[
\tilde{T} = T_{\text{exact}} + \sigma v \tag{8}
\]

where \(\sigma\) is the standard deviation of the errors and \(v\) is a random variable taken from a statistical distribution, with zero mean and unitary variance. All tests were carried out using 5% of noise \((\sigma=0.05)\).

It is important to observe that the spatial grid consists of 101 points \(N_\text{x}=101\), and the time-integration is performed up to \(\tau=0.01\). If we effectively want to apply some kind of regularization, which means \(\alpha>0\) in Eq. (5), then
the discrepancy principle - an a-posteriori parameter choice rule - implies that a suitable regularized solution can be obtained. Since the spatial resolution is \( N_x = 101 \), the optimum \( \alpha \) is reached for \( R(f^*) = N_x \sigma^2 = 0.2525 \) (according to the condition (6)).

The parameter vector was always subjected to the following simple bounds: \( 1.2 \geq f_k \geq -0.2 \) for the triangular test function, and \( 1.2 \geq f_k \geq 0 \) for the semi-triangular test function, with \( k = 1, 2, \ldots, N_x \).

**Probability Density Functions (PDF)**

For generating the random variable in Eq. (8), several distributions have been considered.

- **Uniform distribution:**
  \[
  \rho(x) = u(x; c, d) = \begin{cases} 1/(c-d) & \text{for } c \leq x \leq d \\ 0 & \text{otherwise} \end{cases}
  \]  
  (9)

- **Normal (Gaussian) distribution:**
  \[
  \rho(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2}
  \]  
  (10)

where \( \mu \) and \( \sigma \) are mean and standard deviation, respectively.

- **Cauchy’s distribution:**
  \[
  \rho(x) = \frac{s}{\pi ((x-t)^2 + s^2)}
  \]  
  (11)

where \( t \) is the location parameter and \( s \) is the scale parameter. The case where \( t=0 \) and \( s=1 \) is called the standard Cauchy distribution, and the PDF reduces to

\[
\rho(x) = \frac{1}{\pi (1 + x^2)}
\]  
(12)

- **Student’s t distribution:**
  \[
  \rho(x) = \frac{\Gamma\left(\frac{a+1}{2}\right)}{\sqrt{\pi a} \Gamma\left(\frac{a}{2}\right)}\left(1 + \frac{x^2}{a}\right)^{-\frac{a+1}{2}}
  \]  
  (13)

where \( a \) is a parameter, and \( \Gamma \) is the Gamma function.

- **Tsallis’s distribution:**
  A non-extensive form of entropy has been proposed by Tsallis [15]:

\[
S_q(p) = \frac{k}{q-1} \left(1 - \sum_{i=1}^{N_x} p_i^q\right)
\]  
(14)

where \( p_i \) is a probability, and \( q \) is a free parameter - it is called the non-extensivity parameter, and the parameter \( q \) has a central role in Tsallis’ thermostatistics. In thermodynamics the parameter \( k \) is known as the Boltzmann’s constant. Tsallis’ entropy reduces to the the usual Boltzmann-Gibbs-Shanon formula

\[
S(p) = -k \sum_{i=1}^{N_x} p_i \ln p_i
\]  
(15)

in the limit \( q \to 1 \).

As for extensive form of entropy, the equiprobability condition produces the maximum for the non-extensive entropy function, and this condition leads to special distributions:

\( q > 0 \):

\[
\rho(x) = \frac{1}{\sigma} \left[ \frac{q-1}{\pi (3-q)} \right]^{1/2} \left( \frac{1}{\Gamma\left(\frac{3-q}{2}\right)} \right) \left( \frac{1}{\Gamma\left(\frac{1}{q-1}\right)} \right) \frac{1}{2(q-1)\Gamma\left(\frac{2}{3-q} \sigma^2\right)}
\]  
(16)

\( q = 0 \):

\[
\rho(x) = \frac{1}{\sigma} \left[ \frac{1}{2\pi} \right]^{1/2} e^{-\left(x/\sigma\right)^2/2}
\]  
(17)

\( q < 0 \):

\[
\rho(x) = \frac{1}{\sigma} \left[ \frac{1}{\pi (3-q)} \right]^{1/2} \left( \frac{1}{\Gamma\left(\frac{5-3q}{2(1-q)}\right)} \right) \left( \frac{1}{\Gamma\left(\frac{2-q}{(3-q)\sigma^2}\right)} \right) e^{-\left(x/\sigma\right)^2/(1-q)}
\]  
(18)
if \( x < \sigma[(3-q)/(1-q)]^{1/2} \), \( p(x) = 0 \) otherwise. These PDFs are shown in Figure 1.

For \( q < 5/3 \), the standard central limit theorem applies, implying that if \( p_i \) is written as a sum of \( M \) random independent variables, in the limit case \( M \to \infty \), the probability density function for \( p_i \) in the distribution space is the normal (Gaussian) distribution [6]. However, for \( 5/3 < q < 3 \) the Levy-Gnedenko’s central limit theorem applies, resulting for \( M \to \infty \) the Levy distribution as the probability density function for the random variable \( p_i \). The index in such Levy distribution is \( \gamma = (3-q)/(q-1) \) [16].

These PDFs are shown in Figure 1.

For \( q < 5/3 \), the standard central limit theorem applies, implying that if \( p_i \) is written as a sum of \( M \) random independent variables, in the limit case \( M \to \infty \), the probability density function for \( p_i \) in the distribution space is the normal (Gaussian) distribution [6]. However, for \( 5/3 < q < 3 \) the Levy-Gnedenko’s central limit theorem applies, resulting for \( M \to \infty \) the Levy distribution as the probability density function for the random variable \( p_i \). The index in such Levy distribution is \( \gamma = (3-q)/(q-1) \) [16].

**Generalized Discrepancy Principle**

Should be noted that Cauchy’s distribution (11), Student’s \( t \) distribution, and Tsallis’s distribution (with \( q > 5/3 \)), their second statistical moments diverge (\( \sigma^2 \to \infty \)).

Firstly, all distributions are normalized (when this is possible) for \( \sigma^2 = 1 \). For distributions with divergent variance, a modified PDF is considered, in such a way that a parameter \( d \) is chosen for satisfying the relation

\[
\int_{-d}^{d} x^2 p(x) dx = 1. \tag{19}
\]

The PDF and this new domain \([-d, d]\) is now used for generating the random variable \( \nu \) in Eq. (8), and the discrepancy principle can be applied for any distribution type.

**Numerical Results**

The numerical experiments were carried out for many realizations (10). For each experiment, new different random numbers were generated. For all cases the deviation was assumed as \( \sigma = 0.05 \), as mentioned before, and the discrepancy principle (6) was used to compute the regularization parameter.

Tables 1 and 2 show computed values for regularization parameter using the generalized discrepancy principle. Inverse solutions are depicted in figures below.

**Table 1. Triangular test function**

<table>
<thead>
<tr>
<th>Distribution</th>
<th>( \tilde{\alpha} )</th>
<th>( \tilde{E}(f_u) )</th>
<th>( \sigma E(f_u) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uniform</td>
<td>722.57</td>
<td>0.3929</td>
<td>0.0396</td>
</tr>
<tr>
<td>Normal</td>
<td>719.42</td>
<td>0.3738</td>
<td>0.0327</td>
</tr>
<tr>
<td>Cauchy</td>
<td>685.69</td>
<td>0.3956</td>
<td>0.0499</td>
</tr>
<tr>
<td>Student’s ( t )</td>
<td>719.10</td>
<td>0.4112</td>
<td>0.0576</td>
</tr>
<tr>
<td>Tsallis (( q=1.5 ))</td>
<td>648.49</td>
<td>0.3589</td>
<td>0.0594</td>
</tr>
<tr>
<td>Tsallis (( q=0.5 ))</td>
<td>767.34</td>
<td>0.4268</td>
<td>0.0470</td>
</tr>
</tbody>
</table>

**Table 2. Semi-triangular test function**

<table>
<thead>
<tr>
<th>Distribution</th>
<th>( \tilde{\alpha} )</th>
<th>( \tilde{E}(f_u) )</th>
<th>( \sigma E(f_u) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uniform</td>
<td>682.01</td>
<td>0.4698</td>
<td>0.0895</td>
</tr>
<tr>
<td>Normal</td>
<td>676.58</td>
<td>0.4571</td>
<td>0.0509</td>
</tr>
<tr>
<td>Cauchy</td>
<td>632.44</td>
<td>0.4777</td>
<td>0.1045</td>
</tr>
<tr>
<td>Student’s ( t )</td>
<td>677.19</td>
<td>0.5311</td>
<td>0.1197</td>
</tr>
<tr>
<td>Tsallis (( q=1.5 ))</td>
<td>601.23</td>
<td>0.4580</td>
<td>0.0786</td>
</tr>
<tr>
<td>Tsallis (( q=0.5 ))</td>
<td>737.94</td>
<td>0.5383</td>
<td>0.0726</td>
</tr>
</tbody>
</table>
- Normal (Gaussian) distribution:

- Cauchy’s distribution:

- Tsallis’s distribution:

Figure 2. Initial condition estimation, with uniform distribution noise.

Figure 3. Initial condition estimation, with normal distribution noise.

Figure 4. Initial condition estimation, with Cauchy distribution noise.
CONCLUSION

It was shown that the discrepancy principle can be adapted for situations where the error random variable can follow other statistical distributions instead of Gaussian distribution. Even those distributions that do not have a defined variance, the PDF of the random number generator can be modified, becoming possible to apply the discrepancy principle. For the worked examples, the regularization parameter can be determined as exposed, producing good inverse solutions.

ACKNOWLEDGEMENTS

The authors thanks to the FAPESP and CNPq, Brazilian agencies for research financial support, and an anonymous referee, whose comments improved this version of the paper.

REFERENCES

9. L.D. Chiwiacowsky, H.F. de Campos Velho, Different Approaches for the Solution of a


