

## A unified regularization theory: the maximum non-extensive entropy principle

HAROLDO F. DE CAMPOS VELHO<sup>1</sup>, ELCIO H. SHIGUEMORI<sup>2</sup>  
FERNANDO M. RAMOS<sup>1</sup> and JOÃO C. CARVALHO<sup>3</sup>

<sup>1</sup>Instituto Nacional de Pesquisas Espaciais (INPE)  
Laboratório Associado de Computação e Matemática Aplicada (LAC)  
São José dos Campos, SP, Brazil

<sup>2</sup>Instituto de Estudos Avançados (IEAv)  
Comando Geral Brasileiro de Tecnologia Aeroespacial (CTA)  
São José dos Campos, SP, Brazil

<sup>3</sup>Agência Nacional de Telecomunicações (ANATEL)  
Brasília, DF, Brazil

E-mails: [haroldo, elcio, fernando]@lac.inpe.br / jcarvalho@anatel.gov.br

---

**Abstract.** Tsallis' non-extensive entropy is used as a regularization operator. The parameter “ $q$ ” (non-extensivity parameter) has a central role in the Tsallis' thermostatics formalism. Here, several values of  $q$  are investigated in inverse problems, using  $q < 1$  and  $q > 1$ . Two standard regularization techniques are recovered for special  $q$ -values: (i)  $q = 2$  is the well known Tikhonov regularization; (ii)  $q = 1$  is the standard Boltzmann-Gibbs-Shannon formulation for entropy. The regularization feature is illustrated in an inverse test problem: the estimation of initial condition in heat conduction problem. Two methods are studied for determining the regularization parameter, the maximum curvature for the  $L$ -curve, and the Morozov's discrepancy principle. The new regularization of higher order is applied to the retrieval of the atmospheric vertical profile from satellite data.

**Mathematical subject classification:** 35R30.

**Key words:** Maximum non-extensive entropy principle, unified regularization theory, inverse problems.

## 1 Introduction

Inverse problems belong to the class of *ill*-posed problems. In these type of problems existence, uniqueness and stability of their solutions cannot be ensured.

The regularization theory appeared in the 60's as a general method for solving inverse problems [1]. In this approach, the non-linear least square problem is associated with a regularization term (*a priori* or additional information), in order to obtain a well-posed problem. A well-known regularization technique was proposed by Tikhonov [1], where smoothness of the unknown function is searched. Similarly to Tikhonov's regularization, the maximum entropy formalism searches for *global* regularity and yields the smoothest reconstructions which are consistent with the available data. The maximum entropy principle was first proposed as a general inference procedure by Jaynes [2] on the basis of Shannon's axiomatic characterization of the amount of information [3]. This principle has successfully been applied to a variety of fields. In the middle of 90's, higher order procedures for entropic regularization have been introduced [5, 6] (see also [7] and references inside).

In the macroscopic thermodynamics, the entropy is applied to quantify the irreversibility in systems, and the entropy has a fundamental role in the statistical physics. The Shannon's entropy is a measurement of the missing information. In the inverse problem context, it is a *shape entropy*.

A non-extensive formulation for the entropy has been proposed by Tsallis [8, 9]. Recently, the non-extensive entropic form ( $S_q$ ) was used as a new regularization operator [10], using only  $q = 0.5$ . The  $q$  parameter plays a central role in the Tsallis' thermostatistics, in which  $q = 1$  the Boltzmann-Gibbs-Shannon's entropy is recovered. As mentioned, the non-extensive entropy includes as a particular case the extensive entropy ( $q = 1$ ): in which the maximum entropy principle can be used as a regularization method. However, another important particular case of the non-extensive entropy occurs when  $q = 2$ . In such case, the maximum *non-extensive* entropy principle to the  $S_2$  regularization operator is equivalent to the standard Tikhonov regularization (or zeroth-order Tikhonov regularization) [1, 11].

This new regularization operator was tested for estimating initial condition in heat conduction problem [11, 12]. In the present study several values for

$q$  were used for the *non-extensive* entropic regularization term. Synthetic data with Gaussian white noise corruption were used to simulate experimental data. Two methods were investigated for determining the regularization parameter: the Morozov’s discrepancy principle [13], and the maximum curvature scheme of the curve relating smoothness versus fidelity, inspired in Hansen’s geometrical criterion [14].

Initial conditions for numerical weather prediction are obtained from the Earth observation system. Radiances measured by meteorological satellites is a component of this observation system. From these radiances, atmospheric temperature and humidity profiles can be determined. Several methodologies and models have been developed to compute this estimation. A method using the second order entropy approach [7] was proposed as a solution for this inverse problem. Here, the higher order non-extensive entropy will be used to improve the previous inverse solution.

**Non-extensive entropy as a new regularization operator**

A non-extensive form of entropy has been proposed by Tsallis [8] by the expression

$$S_q(\mathbf{p}) = \frac{k}{q-1} \left[ 1 - \sum_{i=1}^{N_p} p_i^q \right] \tag{1}$$

where  $p_i$  is a probability, and  $q$  is a free parameter. In thermodynamics the parameter  $k$  is known as the Boltzmann’s constant. Similarly as in the mathematical theory of information,  $k = 1$  is considered in the regularization theory. Tsallis’ entropy reduces to the the usual Boltzmann-Gibbs-Shanon formula

$$S_q(\mathbf{p}) = -k \sum_{i=1}^{N_p} p_i \ln p_i \tag{2}$$

in the limit  $q \rightarrow 1$ .

As for extensive form of entropy, the equiprobability condition produces the maximum for the non-extensive entropy function, and this condition is expressed as

$$(S_q)_{\max} = \frac{1}{1-q} (N_p^{1-q} - 1) \quad (N_p \geq 1) \tag{3}$$

where, in the limit  $N_p \rightarrow \infty$ ,  $S_q$  diverges if  $q \leq 1$ , and saturates at  $1/(q - 1)$  if  $q > 1$  [9].

The equiprobability condition leads the regularization function defined by the operator  $S_q(\mathbf{p})$ , given by Eq. (1), to search the smoothest solution for the of the unknown vector  $\mathbf{p}$ .

The parameter  $q$  has a central role in Tsallis' thermostats, and it is called the *non-extensivity parameter* – see Eqs. (A.5) and (A.6) in the Appendix. Figure 1 shows the functional form for the Tsallis' entropy for several values of  $q$ . 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  $M \rightarrow \infty$ , the probability density function for  $p_i$  in the distribution space is the *normal* (Gaussian) distribution [9]. However, for  $5/3 < q < 3$  the Levy-Gnedenko's central limit theorem applies, resulting for  $M \rightarrow \infty$  the Levy distribution as the probability density function for the random variable  $p_i$ . The index in such Levy distribution is  $\gamma = (3 - q)/(q - 1)$  [9].

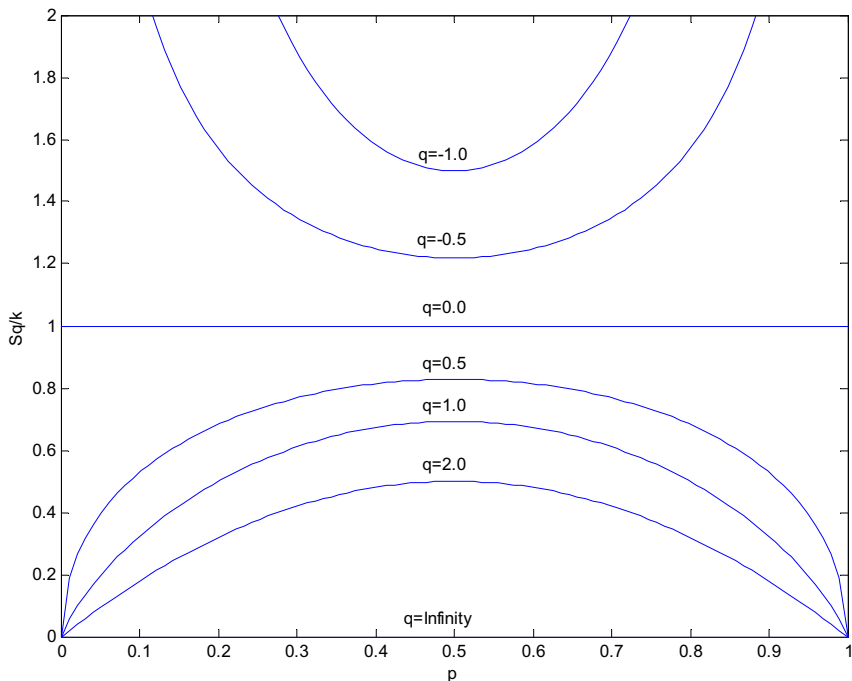


Figure 1 – The behaviour of the non-extensive entropy function for several values of  $q$ .

The non-extensive approach has been used in many different applications, such as in a certain type of anomalous diffusion process [9], as well as the statistical model for data from turbulent flow [15] and from financial market [16]. According to Plastino and Plastino [17], the first experimental confirmation of Tsallis' non-extensive formalism is the Boghosian's approach of the two dimensional pure electron plasma [18]. Some properties of the thermostatistics formalism are described in the Appendix.

*Regularization unified*

Some comments have been addressed for indicating the physical properties of the  $S_q$  operator. The goal of this Section is to describe formally the properties for this operator looking at regularization purposes. Regularization properties for entropy operator emerges from the Jaynes' inference criterium: the maximum entropy principle, where all events have the same propability to occur. Implying that all parameters assume the same value:  $p_i = 1/N_p$ . The following Lemma extends this result for non-extensive entropy.

**Lemma.** *The non-extensive function  $S_q$  is maximum as  $p_i = 1/N_p$  for all  $i$ .*

**Proof.** The problem is to find the maximum of the function (1), with the following constrain

$$\sum_{i=1}^{N_p} p_i = 1 \tag{4}$$

since  $p_i$  represents a probability. Therefore, it is possible to define an objective function where the constrain can be added to the non-extensive function:

$$J(\mathbf{p}) = S_q(\mathbf{p}) + \lambda \left( \sum_{i=1}^{N_p} p_i - 1 \right) \tag{5}$$

where  $\lambda$  is the Langrange multiplier. The Lagrange multiplier, in this case, can be determined when a minimum for the objective function  $J(\mathbf{p})$  is found, as following

$$\frac{\partial J}{\partial p_i} = -q p_i^{q-1} + \lambda(q - 1) = 0 \Rightarrow p_i = \left[ \frac{\lambda(q - 1)}{q} \right]^{\frac{1}{q-1}} \tag{6}$$

This result can be used to obtain the value of the  $p_i$ 's that maximizes the function  $J(\mathbf{p})$ :

$$\sum_{i=1}^{N_p} p_i = \sum_{i=1}^{N_p} \left[ \frac{\lambda(q-1)}{q} \right]^{\frac{1}{q-1}} = \left[ \frac{\lambda(q-1)}{q} \right]^{\frac{1}{q-1}} N_p = 1 \Rightarrow p_i = \frac{1}{N_p} \quad (7)$$

that means if  $p_i = 1/N_p$  for all  $i = 1, \dots, N_p$  the non-extensive entropy function is maximum.

The next theorem shows that the extensive entropy and Thikhonov's regularizations are particular cases of the non-extensive entropy.

**Theorem.** For particular values for non-extensive entropy  $q = 1$  and  $q = 2$  are equivalents to the extensive entropy and Tikhonov regularizations, respectively.

**Proof.** (i)  $q = 1$ : Taking the limit,

$$\begin{aligned} \lim_{q \rightarrow 1} S_q(\mathbf{p}) &= \lim_{q \rightarrow 1} \frac{1 - \sum_{i=1}^{N_p} p_i^q}{q-1} = \lim_{q \rightarrow 1} \frac{1 - \sum_{i=1}^{N_p} e^{q \log p_i}}{q-1} \\ &= \lim_{q \rightarrow 1} \frac{-\sum_{i=1}^{N_p} \log p_i e^{q \log p_i}}{1} = -\sum_{i=1}^{N_p} p_i \log p_i \end{aligned} \quad (8)$$

(ii)  $q = 2$ : Remembering that  $\max S_2$  is equivalent to  $\min(-S_2)$ , yields

$$\begin{aligned} \max S_2(\mathbf{p}) &= \max \left( 1 - \sum_{i=1}^{N_p} p_i^2 \right) \Leftrightarrow \min [-S_2(p)] \\ &= \min \left( \sum_{i=1}^{N_p} p_i^2 - 1 \right) \end{aligned} \quad (9)$$

now, for the maximum (minimum) value holds  $\nabla_{\mathbf{p}} S_2 = 0$ , therefore

$$\nabla_{\mathbf{p}} S_2(\mathbf{p}) = \nabla_{\mathbf{p}} \left( \sum_{i=1}^{N_p} p_i^2 - 1 \right) = \nabla_{\mathbf{p}} \left( \sum_{i=1}^{N_p} p_i^2 \right) = \nabla_{\mathbf{p}} \|\mathbf{p}\|_2^2 \quad (10)$$

In conclusion:  $\max S_2(\mathbf{p}) = \min \|\mathbf{p}\|_2^2$  (the zeroth-order Tikhonov regularization).

## Inverse analysis

Typically, inverse problems are ill-posed – existence, uniqueness and stability of their solutions cannot be ensured. An inverse solution can be formulated to obtain existence and uniqueness, but this solution can still be unstable under the presence of noise in the experimental data. Hence, it requires some regularization technique, i.e., the incorporation in the inversion procedure of some available information about the true solution. Following the Tikhonov's approach [1], a regularized solution is obtained by choosing the function  $p^*$  that minimizes the following functional

$$J_\alpha[\tilde{\Phi}, p] = \left\| \tilde{\Phi} - \Phi(p) \right\|_2^2 + \alpha \Omega[p] \quad (11)$$

where  $\tilde{\Phi}$  is the experimental data,  $\Phi(p)$  is the answer computed from the forward model,  $\Omega[p]$  denotes the regularization operator,  $\alpha$  is the regularization parameter, and  $\|\cdot\|_2$  is the  $L_2$  norm.

The regularization parameter  $\alpha$  is chosen by two methods: numerically, assuming that a bound  $\delta$  (or the 'statistics') of the measurement error is known, i.e.,  $\left\| \Phi_{\text{computed}} - \tilde{\Phi} \right\| \leq \delta$  — this numerical procedure is based on Morozov's discrepancy principle [13]; graphically, finding out the point of maximum curvature in the curve  $\Omega[p_\alpha] \times \left\| \tilde{\Phi} - \Phi(p_\alpha) \right\|_2^2$ , a type of L-curve [14, 19].

### *Optimization algorithm*

The optimization problem is iteratively solved by the quasi-newtonian optimizer routine from the NAG Fortran Library [20]. This algorithm is designed to minimize an arbitrary smooth function subject to constraints (simple bound, linear or non-linear constraints), using a sequential programming method.

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

## Backward heat conduction

An example of the use for this new generalized regularization is the identification of the initial condition in heat conduction transfer. 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

$$\frac{\partial^2 T(x, t)}{\partial x^2} = \frac{\partial T(x, t)}{\partial t}, \quad x \in (0, 1), \quad t > 0, \quad (12)$$

$$\frac{\partial T(x, t)}{\partial x} = 0, \quad x = 0; \quad x = 1, \quad t > 0, \quad (13)$$

$$T(x, t) = f(x), \quad x \in [0, 1], \quad t = 0, \quad (14)$$

where  $T(x, t)$  (temperature),  $f(x)$  (initial condition),  $x$  (spatial variable) and  $t$  (time variable) are dimensionless quantities. 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  $O(\Delta t)$  [21].

This problem has been used for testing different methodologies in inverse problems [12, 11, 22, 23, 24], and it is badly conditioned problem [12].

The numerical experiment with the non-extensive entropy is based on two test functions, the triangular function

$$f(x) = \begin{cases} 2x, & 0 \leq x < 0.5, \\ 2(1-x), & 0.5 \leq x \leq 1; \end{cases} \quad (15)$$

and semi-triangular function

$$f(x) = \begin{cases} 0.55, & 0 \leq x < 0.2, \\ \frac{8}{3}x + \frac{7}{15}, & 0.2 \leq x < 0.5, \\ -\frac{28}{5}x + \frac{23}{5}, & 0.5 \leq x < 0.75, \\ \frac{2}{9}, & 0.75 \leq x \leq 1. \end{cases} \quad (16)$$

The experimental data (measured temperatures at a time  $t > 0$ ), which intrinsically contains errors, is obtained by adding a random perturbation to the exact solution of the direct problem, such that

$$\tilde{T} = T_{\text{exact}} + \sigma \mu \quad (17)$$



where  $\sigma$  is the standard deviation of the errors, and  $\mu$  is a random variable taken from a Gaussian 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 ( $N_x = 101$ ), and the time-integration is performed up to  $t = 0.01$ . The residue  $R(f_\alpha)$  and the error  $E(f_\alpha)$  are defined by

$$R(f_\alpha) \equiv \left\| \tilde{T} - T(f_\alpha) \right\|_2^2; \tag{18}$$

$$E(f_\alpha) \equiv \left\| f_\alpha - f_{\text{exact}} \right\|_2^2. \tag{19}$$

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  $R(f_*) \simeq N_x \sigma^2 = 0.2525$ . Table 1 shows the square difference term  $R(f_*)$  obtained for different values of  $\alpha$ , and the *optimum* value is pointed out for each case in bold font.

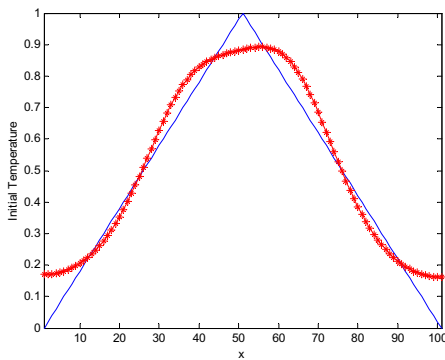
Triangular function			Semi-triangular function		
$\alpha$	$R(f_\alpha)$	$E(f_\alpha)$	$\alpha$	$R(f_\alpha)$	$E(f_\alpha)$
0.0001	0.1853	2.7298	0.0001	0.1851	4.9359
0.0003	0.1856	0.5388	0.0003	0.1854	0.6740
0.0010	0.1861	0.3443	0.0010	0.1856	0.3443
<b>0.0285</b>	<b>0.2525</b>	<b>0.3994</b>	<b>0.0346</b>	<b>0.2525</b>	<b>0.2400</b>
0.0999	0.7728	0.8684	0.0999	0.6687	0.7207

Table 1 – Determining regularization parameter by Morozov’s criterion:  $q = 0.5$ .

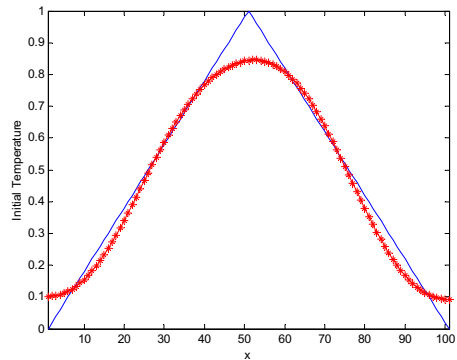
A set of tables (Tables 1 and 2) presents the least squares (or residual) term  $R(f_\alpha)$  and the error  $E(f_\alpha)$  between the approximated (or calculated) solution  $f_\alpha$  and the exact solution  $f_{\text{exact}}$  obtained for two values of  $q = 0.5$  and  $q = 2.0$ , from a family of regularization non-extensive entropy functions  $S_q$ , for different values of  $\alpha$ . The value of  $\alpha$  satisfying the discrepancy principle is pointed out for each test functions by bold font. Regularized solutions are presented in Figures 2 and 3.

Triangular function			Semi-triangular function		
$\alpha$	$R(f_\alpha)$	$E(f_\alpha)$	$\alpha$	$R(f_\alpha)$	$E(f_\alpha)$
0.0001	0.1852	2.3979	0.0001	0.1851	4.0920
0.0003	0.1854	0.3578	0.0003	0.1854	0.4455
0.0010	0.1856	0.1383	0.0010	0.1856	0.1807
<b>0.0414</b>	<b>0.2525</b>	<b>0.2370</b>	<b>0.0419</b>	<b>0.2525</b>	<b>0.4379</b>
0.0999	0.5174	0.5977	0.0999	0.5131	0.8651

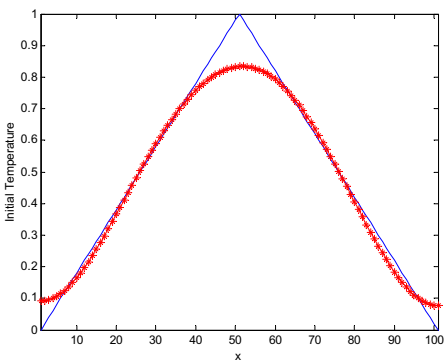
Table 2 – Determining regularization parameter by Morozov’s criterion:  $q = 2.0$ .



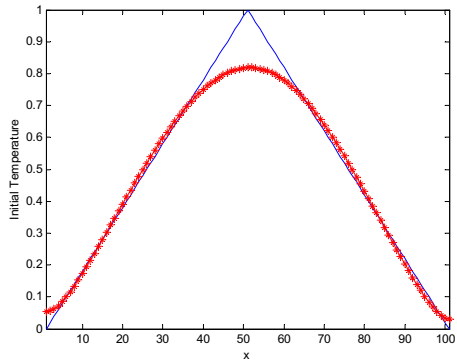
(a)



(b)



(c)



(d)

Figure 2 – Reconstructions with 5% of noise, with a determined by Morozov’s principle: (a)  $q = 0.5$ ; (b)  $q = 1.5$ ; (c)  $q = 2.0$ ; (d)  $q = 2.5$ .

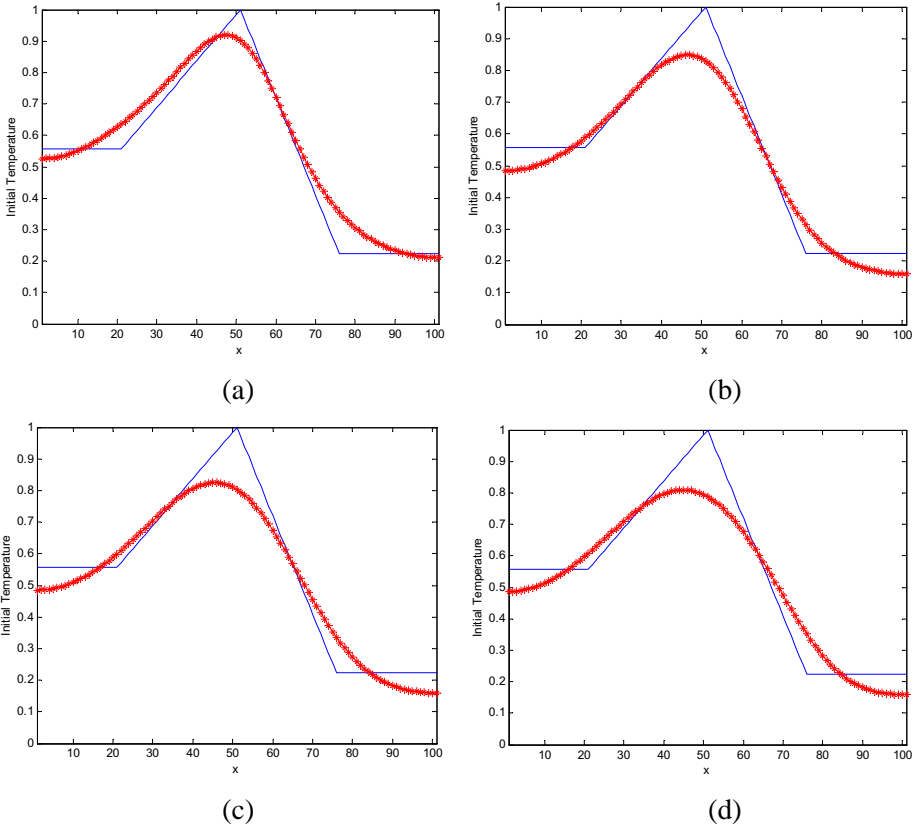


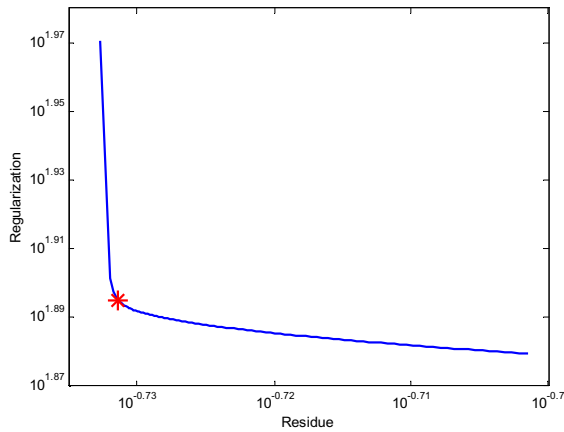
Figure 3 – Reconstructions with 5% of noise, with a determined by Morozov’s principle: (a)  $q = 0.5$ ; (b)  $q = 1.5$ ; (c)  $q = 2.0$ ; (d)  $q = 2.5$ .

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  $k = 1, 2, \dots, N_x$ .

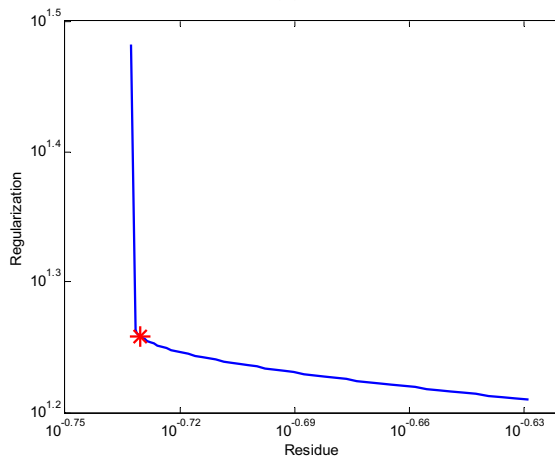
Figures 3a–3d show the estimation of four  $q$ -values for triangular initial condition, where the regularization parameter was computed by Morozov’s principle. The best reconstruction was found by  $q = 2.5$ , but good reconstructions were obtained for other values of  $q$  too. Figures 2a–2d depict the reconstructions for the semi-triangular test function, showing good reconstructions for all values of  $q$ .

Another criterion for finding the regularization parameter was also investi-

gated, and it is based on the maximum curvature in the L-curve [14]. Figures 4a–4b show the L-curve for triangular test function using  $q = 0.5$  and  $2.0$ , respectively. The L-curve for semi-triangular test function is displayed in Figures 5a–5b. The regularization parameter  $\alpha$  is chosen at the corner of the L-curve.



(a)

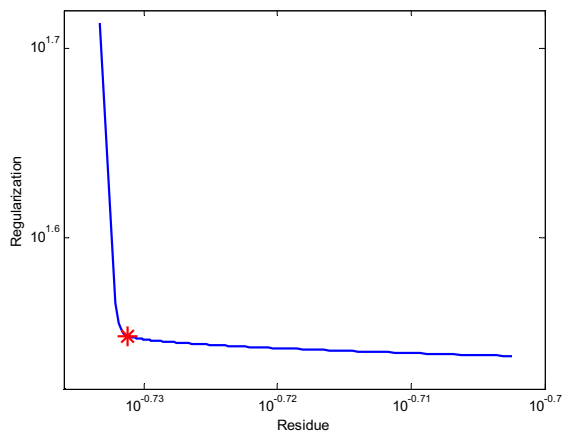


(b)

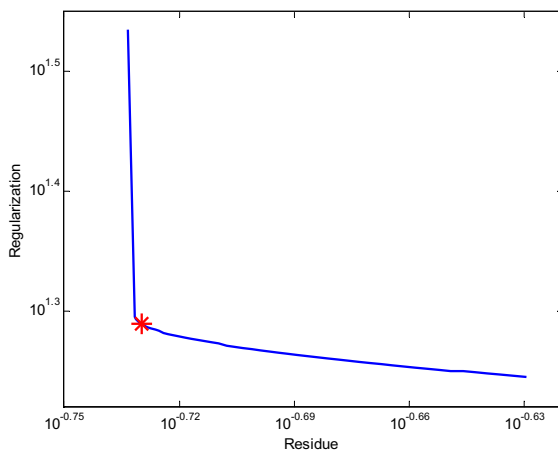
Figure 4 – L-curve for triangular test function: (a)  $q = 1.5$ ; (b)  $q = 2.5$ .

The numerical values for the regularization parameters estimated by Morozov's and Hansen's criteria are shown in Table 3.

Reconstructions using  $\alpha$  as computed by Hansen's criterion [14] are shown in Figures 6a–6d for the triangular test function, and Figures 7a–7d for the semi-



(a)



(b)

Figure 5 – L-curve for semi-triangular test function: (a)  $q = 1.5$ ; (b)  $q = 2.5$ .

triangular test function. The best reconstruction was obtained using  $q = 2.5$ , and the worst for  $q = 0.5$ .

### Retrieval of atmospheric temperature profile

The vertical structure of temperature and water vapor plays an important role in the meteorological process of the atmosphere. However due a several conditions, there is a lack of observation in several regions of the Earth. In this sense, the

	Triangular function		Semi-triangular function	
	$\alpha_{\text{Morozov}}$	$\alpha_{\text{Hansen}}$	$\alpha_{\text{Morozov}}$	$\alpha_{\text{Hansen}}$
$q = 0.5$	0.0285	0.0011	0.0346	0.0040
$q = 1.5$	0.0231	0.0008	0.0231	0.0010
$q = 2.0$	0.0414	0.0040	0.0419	0.0040
$q = 2.5$	0.0579	0.0040	0.0583	0.0050

Table 3 – Regularization parameter computed by test functions using Morozov’s discrepancy principle and Hansen’s criterion.

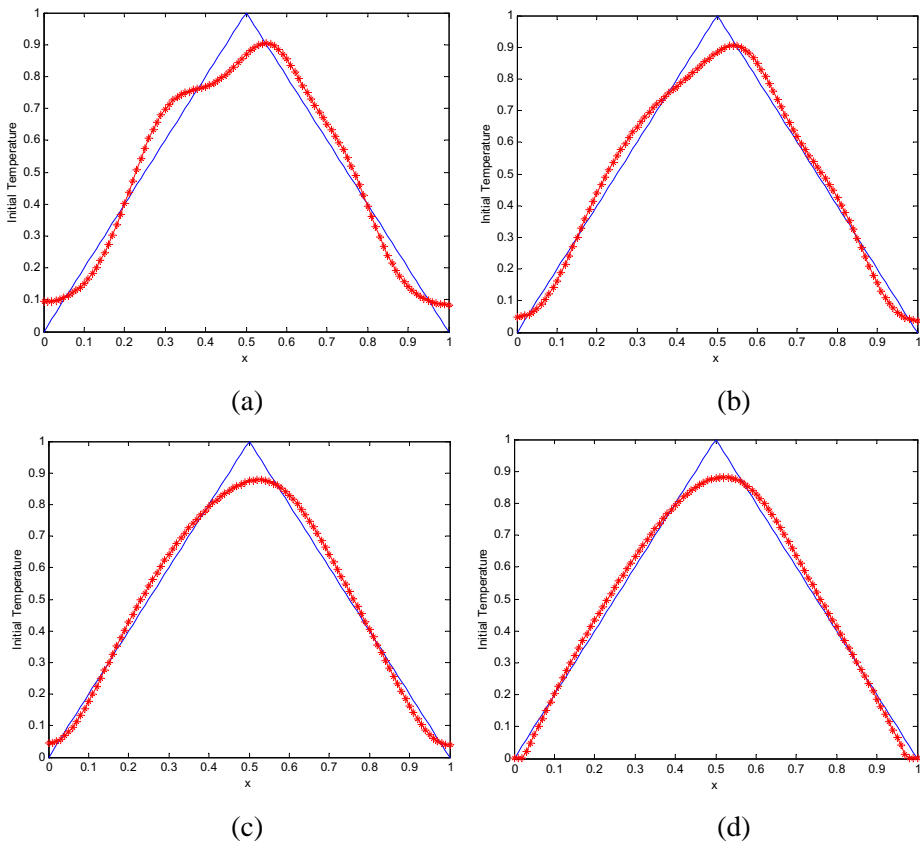


Figure 6 – Reconstructions for triangular test function, with  $\alpha$  determined by Hansen’s criterion: (a)  $q = 0.5$ ; (b)  $q = 1.5$ ; (c)  $q=2.0$ ; (d)  $q=2.5$ .

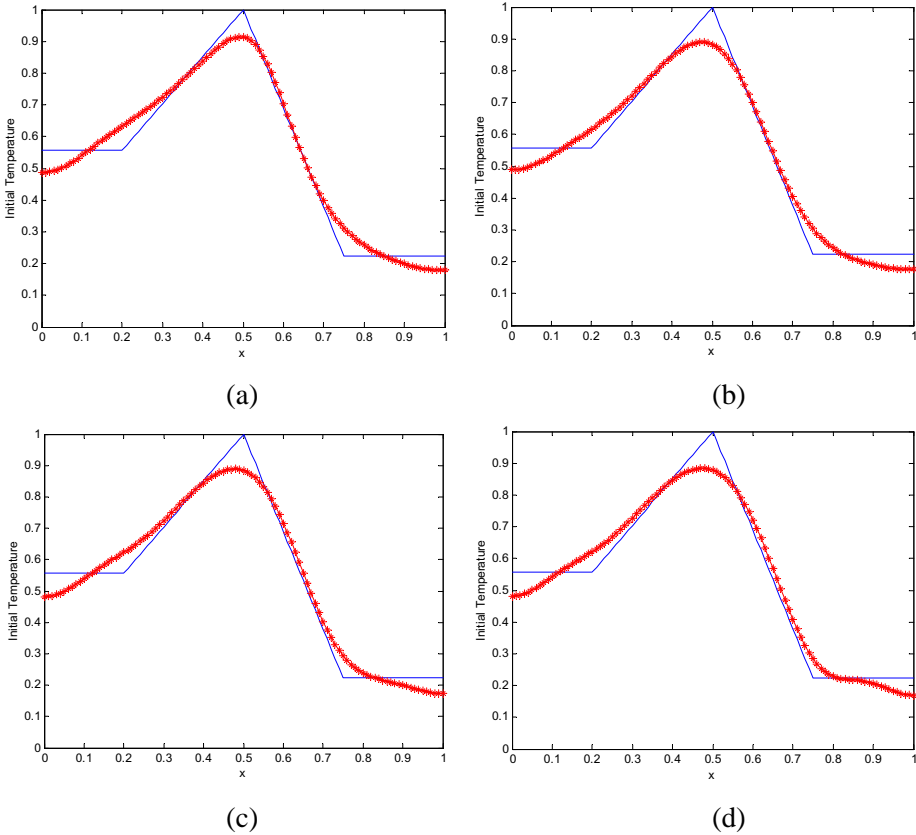


Figure 7 – Reconstructions for semi-triangular test function, with  $\alpha$  determined by Hansen’s criterion: (a)  $q = 0.5$ ; (b)  $q = 1.5$ ; (c)  $q=2.0$ ; (d)  $q=2.5$ .

retrieval of temperature and humidity profiles from satellite radiance data became important for applications such as weather analyses and data assimilation in numerical weather predictions models.

Interpretation of satellite radiances in terms of meteorological fields requires the inversion of the Radiative Transfer Equation (RTE) where measurements of radiation performed in different frequencies are related to the energy from different atmospheric regions. The degree of indetermination is associated with the spectral resolution and the number of spectral channels. Moreover, usually this solution is very unstable regarding the noises in the measuring process. Also, several methodologies and models have been developed to improve the satellite

data processing. Due to the difficulty of obtaining correct RTE solutions, several approaches and methods were developed to extract information from satellite data [25, 26, 27, 28]. The direct problem may be expressed by [29]

$$I_\lambda(0) = B_\lambda(T_s)\mathfrak{S}_\lambda(p_s) + \int_{p_s}^0 B_\lambda[T(p)] \frac{\partial \mathfrak{S}_\lambda(p)}{\partial p} dp, \quad (20)$$

where  $I_\lambda$  is the spectral radiance,  $\lambda$  is the channel frequency;  $\mathfrak{S}$  is the layer to space atmospheric transmittance function, the subscript  $s$  denotes surface [12]; and  $B$  is the Planck function which is a function of the temperature  $T$  (or pressure  $p$ ):

$$B_\lambda(T) = \frac{2hc^2/\lambda^5}{[e^{hc/k_B\lambda T} - 1]} \quad (21)$$

being  $h$  the Planck constant,  $c$  the light speed, and  $k_B$  the Boltzmann constant. For practical purposes, equation (21) is discretized using central finite differences:

$$I_i = B_{i,s}(T_s)\mathfrak{S}_{i,s} + \sum_{j=1}^{N_p} \left( \frac{B_{i,j} + B_{i,j-1}}{2} \right) [\mathfrak{S}_{i,j} - \mathfrak{S}_{i,j-1}] \quad (22)$$

with  $I_i \equiv I_{\lambda_i}(0)$ ,  $N_\lambda$  is the number of channels in the satellite, and  $N_p$  is the number of the atmospheric layers considered.

Some previous results have employed a generalization of the standard maximum entropy principle (MaxEnt-0) for solving this inverse problem [7, 30]: the higher order entropy approach. The same strategy can be applied here. Therefore, the non-extensive entropy of order- $\gamma$  is defined as

$$S_q^\gamma \equiv \frac{k}{q-1} \left[ 1 - \sum_{i=1}^{N_p} r_i^q \right]; \quad r_i = \frac{T_i}{\sum_{i=1}^{N_p} T_i} \quad (23)$$

and

$$\mathbf{r} = \Delta^\gamma \mathbf{p} \quad (24)$$

where  $\gamma = 0, 1, 2, \dots$ , and  $\Delta$  is a discrete difference operator. The standard MaxEnt-0 can be derived from (23) and (24) imposing  $\gamma = 0$  and  $q = 1$ . A small value should be added to the difference operator (say  $\zeta = 10^{-15}$ ) to assure a definite quantity for all values of  $q$ .



Figures 8–10 present the atmospheric temperature retrieval achieved using radiance data from the High Resolution Radiation Sounder (HIRS-2) of NOAA-14 satellite. HIRS-2 is one of the three sounding instruments of the TIROS Operational Vertical Sounder (TOVS). The results for the non-extensive MaxEnt-0 (for short: NE-MaxEnt-0) did not produce good results – Figure 8. However, results with NE-MaxEnt-1 and NE-MaxEnt-2 are compared to those obtained with MaxEnt-2 (such results have already been analysed against the profile computed by ITPP-5, a TOVS processing package [7] employed by weather service research centers throughout the world), and to *in situ* radiosonde measurements.

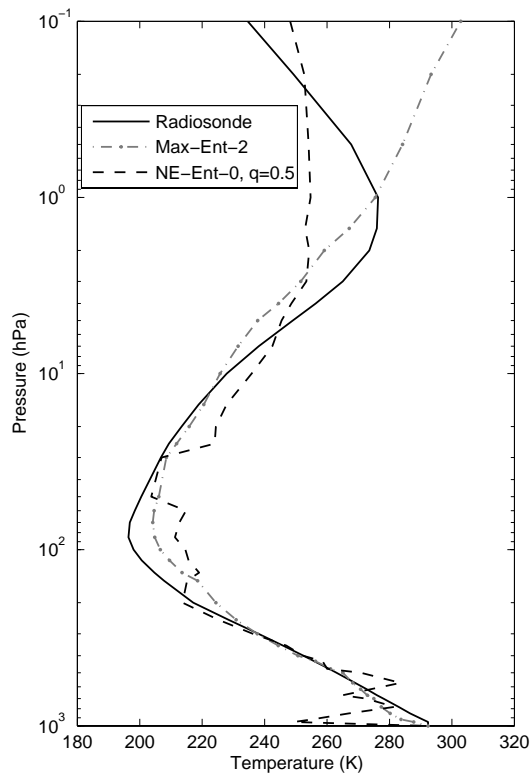


Figure 8 – Reconstructions for temperature profile:  $q = 0.0$ .

From figures 9 and 10 is hard to identify the better performance among the retrievals. Table 5 compares the RMS of NE-MaxEnt-2 and NE-MaxEnt-1 related to the MaxEnt-2. The layer defined by 500–250 hPa there is no significant differences among the different regularization operators. But, for the layer

250–50 hPa, the best performance was obtained by the NE-MaxEnt-1, for the 700–500 hPa better result for the MaxEnt-0, and for the region 1000–750 hPa the NE-MaxEnt-2 ( $q = 0.5$ ) has presented the better inversion.

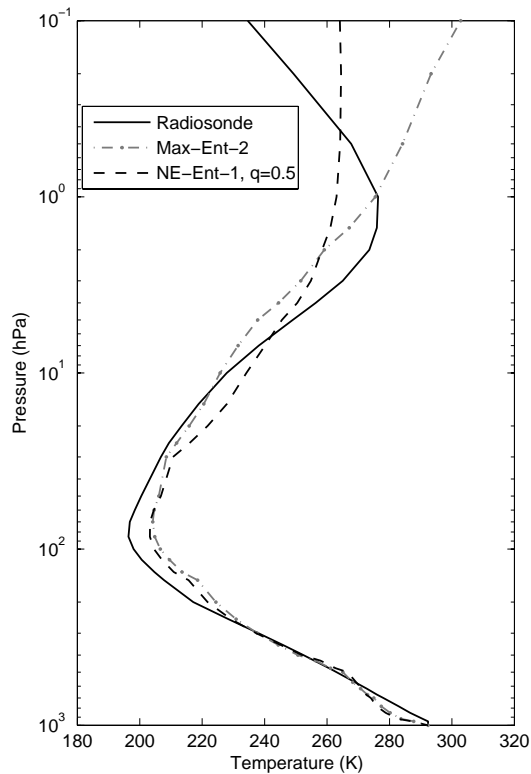


Figure 9 – Reconstructions for temperature profile:  $q = 1.0$ .

## Conclusion

The implicit strategy and the regularization techniques adopted in this work yield good results in reconstructing the initial condition of the heat equation and the atmospheric temperature profile.

For the initial condition estimation in the heat transfer process, the Morozov's discrepancy principle was efficient to estimate the regularization parameter in the analyzed cases. The Hansen's criterion also produced good estimates for the regularization parameter  $\alpha$ . According to the Tables 1 and 2, the value of regular-

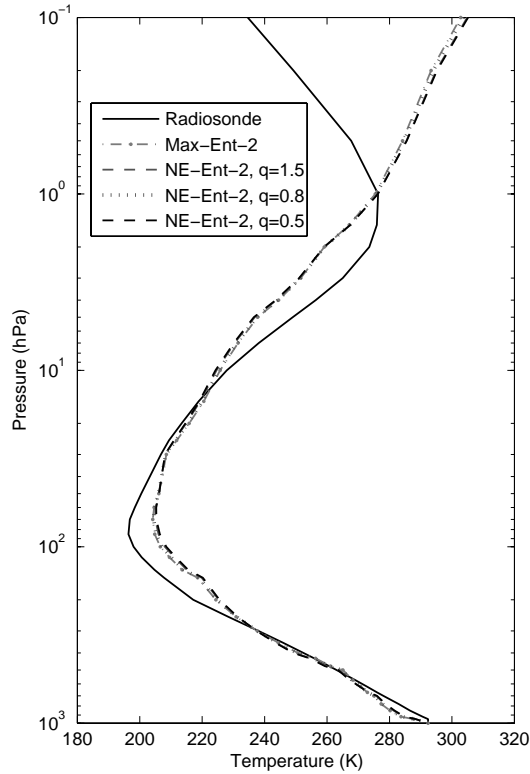


Figure 10 – Reconstructions for temperature profile:  $q = 2.0$ .

ization parameter calculated by the Morozov’s principle tends to over-estimate the value of  $\alpha$ . However, looking at Table 4, it is possible to realize that the parameter  $\alpha$  computed by the Hansen’s criterion is closer to the optimum. Nevertheless, sometimes is hard to obtain the L-curve. One case specially difficult was found to  $q = 2.5$ , for some values of  $\alpha$  was not possible to obtain a solution (no convergence). A possible solution to convergence would be to change the deterministic optimizer by a stochastic one. Table 4 shows that an appropriated choice of the non-extensive parameter  $q$  can improve the reconstruction. Of course, other schemes for determining the regularization parameter can be used (see [31]).

The new regularization technique used in this work, namely the maximum non-extensive entropy, worked very well for the backwards heat equation for all parameter  $q$  tested (zeroth-order). The choice  $q = 0.5$  made by Rebollo-Neira

	Triangular function		Semi-triangular function	
	$E(f_{\alpha\text{-Morozov}})$	$E(f_{\alpha\text{-Hansen}})$	$E(f_{\alpha\text{-Morozov}})$	$E(f_{\alpha\text{-Hansen}})$
$q = 0.5$	0.3994	0.3490	0.2400	0.1676
$q = 1.0$	0.3151	0.1959	0.4056	0.1958
$q = 1.5$	0.2599	0.1697	0.3205	0.1745
$q = 2.0$	0.2370	0.1530	0.4379	0.1800
$q = 2.5$	0.2561	0.1302	0.5959	0.2041

Table 4 – Estimation error computed with regularization parameter found by Morozov's discrepancy principle and Hansen's criterion.

Pressure (hPa)	RMS Max-Ent-2	NE- MaxEnt-2 $q = 1.5$	NE- MaxEnt-2 $q = 0.5$	NE- MaxEnt-1 $q = 0.5$	NE- MaxEnt-0 $q = 0.5$
50-0.1	13.248	13.553	13.646	9.598	10.483
250-50	7.442	8.395	8.723	<b>5.818</b>	11.038
500-250	5.216	5.405	5.508	5.532	10.730
700-500	<b>1.283</b>	1.623	1.817	1.478	2.481
1000-700	4.428	3.704	<b>3.475</b>	5.334	17.518

Table 5 – Root-mean-square for MaxEnt-2 and Non-extensive MaxEnt of first and second orders (NE-MaxEnt-1 and NE-MaxEnt-2).

et al. [10] was linked to a previous result of physical relevance related to the relaxation of two-dimensional turbulence [18]. There is no reason to restrict the regularization operator  $S_q$  at  $q = 0.5$ . Actually, the worse reconstructions for the triangular test function were obtained using  $S_{0.5}$ !

Concerning to the atmospheric temperature retrieval, some improvement was obtained using the higher order of the non-extensive entropic approach. We note that the most important layer for numerical weather prediction lies into the levels 1000-700 hPa. The results suggest that we need to combine different techniques, considering different atmospheric layers, in order to have a better inverse solution.

**A Some properties for the non-extensive thermostatics**

A1: Non-extensive entropy:

$$S_q(p) = \frac{k}{q-1} \left( 1 - \sum_{i=1}^{N_p} p_i^q \right). \tag{25}$$

A2:  $q$ -expectation of an observable:

$$O_q \equiv \langle O \rangle_q = \sum_{i=1}^{N_p} p_i^q o_i. \tag{26}$$

Properties

1. If  $q \rightarrow 1$ :

$$S_1 = k \sum_{i=1}^{N_p} p_i \ln p_i, \tag{27}$$

$$O_1 = \sum_{i=1}^{N_p} p_i o_i. \tag{28}$$

2. Non-extensive entropy is positive:  $S_q \geq 0$ .

3. Non-extensivity

$$S_q(A + B) = S_q(A) + S_q(B) + (1 - q)S_q(A)S_q(B) \tag{29}$$

$$O_q(A + B) = O_q(A) + O_q(B) + (1 - q) [O_q(A)S_q(B) + O_qS_q(A)]. \tag{30}$$

4. Max  $S_q$  under constrain  $O_q = \sum_i p_i^q \epsilon_i$  (canonical ensemble):

$$p_i = \frac{1}{Z_q} [1 - \beta(1 - q)\epsilon_i]^{1/(1-q)} \tag{31}$$

where the  $\epsilon_i$  is the energy of state  $i$ ,  $O_q = U_q$  is the non-extensive form to the internal energy, and the normalization factor  $Z_q$  (partition function), for  $1 < q < 3$ , is given by

$$Z_q = \left[ \frac{\pi}{\beta(1 - q)} \right]^{1/2} \frac{\Gamma[(3 - q)/2(q - 1)]}{\Gamma[1/(q - 1)]}. \tag{32}$$

For  $q = 1$  yields

$$p_i = e^{\beta\epsilon_i} / Z_1. \tag{33}$$

**Acknowledgment.** Authors acknowledge the CNPq and FAPESP, Brazilian agencies for research support. H.F. de Campos Velho also thanks to the Prof. Dr. João Batista da Silva (UFPA, Brazil) and Dra. Valeria Cristina Barbosa (LNCC, Brazil) for the usefull discussions and support.

#### REFERENCES

- [1] A.N. Tikhonov and V.I. Arsenin, *Solutions of Ill-posed Problems*. John Wiley & Sons, 1977.
- [2] E.T. Jaynes, Information theory and statistical mechanics. *Phys. Rev.*, **106** (1057), 620.
- [3] C.D. Rodgers, Retrieval of the atmospheric temperature and composition from remote measurements of thermal radiation. *Rev. Geophys. Space Phys.*, **14** (1976), 609–624.
- [4] C.R. Smith and W.T. Grandy (Eds.). *Maximum-Entropy and Bayesian Methods in Inverse Problems*, in *Fundamental Theories of Physics*, Reidel, Dordrecht (1985).
- [5] F.M. Ramos and H.F. de Campos Velho, Reconstruction of geoelectric conductivity distributions using a minimum first-order entropy technique, 2nd International Conference on Inverse Problems on Engineering, Le Croisic, France, vol. 2, 199 (1996).
- [6] H.F. de Campos Velho and F.M. Ramos, Numerical inversion of two-dimensional geoelectric conductivity distributions from eletromagnetic ground data. *Braz. J. Geophys.*, **15** (1997), 133.
- [7] F.M. Ramos, H.F. de Campos Velho, J.C. Carvalho and N.J. Ferreira, Novel Approaches on Entropic Regularization. *Inverse Problems*, **15**(5) (1999), 1139–1148.
- [8] C. Tsallis, Possible generalization of Boltzmann-Gibbs statistics. *J. Statistical Physics*, **52** (1988), 479.
- [9] C. Tsallis, Nonextensive statistics: theoretical, experimental and computational evidences and connections. *Braz. J. Phys.*, **29** (1999), 1.
- [10] L. Rebollo-Neira, J. Fernandez-Rubio and A. Plastino, A non-extensive maximum entropy based regularization method for bad conditioned inverse problems. *Phys. A*, **261** (1998), 555.
- [11] W.B. Muniz, F.M. Ramos and H.F. de Campos Velho, Entropy- and Tikhonov-based regularization techniques applied to the backwards heat equation. *Comp. Math. Appl.*, **40** (2000), 1071.
- [12] W.B. Muniz, H.F. de Campos Velho and F.M. Ramos, A comparison of some inverse methods for estimating the initial condition of the heat equation. *J. Comp. Appl. Math.*, **103** (1999), 145.
- [13] V.A. Morosov, *Methods for Solving Incorrectly Posed Problems*. Springer Verlag (1984).
- [14] P.C. Hansen, Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Rev.*, **34** (1992), 561.

- [15] F.M. Ramos, R.R. Rosa, C. Rodrigues Neto, M.J.A. Bolzan, L.D.A. Sa and H.F. de Campos Velho, Nonextensive statistics and three-dimensional fully developed turbulence. *Phys. A*, **295** (2001a), 250.
- [16] F.M. Ramos, C. Rodrigues Neto, R.R. Rosa, L.D. Abreu Sa and M.J.A. Bolzan, Generalized thermostistical description of intermittency and nonextensivity in turbulence and financial markets. *Nonlinear Anal.-Theor.*, **23** (2001b), 3521.
- [17] A. Plastino and A.R. Plastino, Tsallis's entropy and Jaynes' information theory formalism. *Braz. J. Phys.*, **29** (1999), 50.
- [18] B.M.R. Boghosian, Thermodynamic description of two-dimensional Euler turbulence using Tsallis' statistics. *Phys. Rev.*, **E 53** (1996), 4754.
- [19] T. Reginska, A regularization parameter in discrete ill-posed problem. *SIAM J. Sci. Comput.*, **17** (1996), 740.
- [20] E04UCF routine, NAG Fortran Library Mark 17, Oxford, UK (1995).
- [21] J.D. Hoffmann, Numerical Methods for Engineers and Scientists. McGraw-Hill, 1993.
- [22] E. Issamoto, F.T. Miki, J.I. da Luz, J.D. da Silva, P.B. de Oliveira and H.F. de Campos Velho, An inverse initial condition problem in heat conduction: a neural network approach, Braz. Cong. on Mechanical Eng. (COBEM), Campinas (SP), Brazil, Proc. in CD-ROM – paper code AAAGHA, page 238 in the Abstract Book (1999).
- [23] F.T. Miki, E. Issamoto, J.I. da Luz, P.B. de Oliveira, H. F. de Campos Velho and J.D. da Silva, A neural network approach in a backward heat conduction problem, Braz. Conference on Neural Networks, Proc. in CD-ROM, paper code 0008, pages 19–24, São José dos Campos (SP), Brazil (1999).
- [24] E.H. Shiguemori, H.F. de Campos Velho and J.D.S. da Silva, Estimation of initial condition in heat conduction by neural network, in this Proceedings – ICIPE-2002, May 26-31, Angra dos Reis (RJ), Brazil.
- [25] C.D. Rodgers, Retrieval of the atmospheric temperature and composition from remote measurements of thermal radiation. *Rev. Geophys. Space Phys.*, **14** (1976), 609–624.
- [26] S. Twomey, Introduction to the mathematics of inversion in remote sensing and iterative measurements. Amsterdam, Elsevier Scientific, 1977.
- [27] W.L. Smith, H.M. Woolf and A.J. Schriener, Simultaneous retrieval os surface and atmospheric parameters: a physical analytically direct approach. *Adv. In Rem. Sens.*, **7** (1985).
- [28] M.T. Chahine, Inverse Problem in Radiative Transfer: determination of atmospheric parameters. *Jour. Atmos. Sci.*, **27** (1970), 960.
- [29] K.N. Liou, An introduction to atmospheric radiation. Academic Press, Orlando, 1982.

- [30] J.C. Carvalho, F.M. Ramos, N.J. Ferreira and H.F. de Campos Velho, Retrieval of Vertical Temperature Profiles in the Atmosphere Inverse Problems in Engineering (3ICIPE), Proceedings in CD-ROM, under paper code HT02 (see also in the internet: <http://www.me.ua.edu/3icipe/fin3prog.htm>) – Proc. Book: pp. 235-238, Port Ludlow, Washington, USA, June 13-18, UEF-ASME (2000).
- [31] M. Bertero and P. Boccacci, Introduction to Inverse Problems in Imaging, Institute of Physics, (1999).