

# Recovering a Probability Density from a Finite Number of Moments and Local a Priori Information

DARIO FASINO and GABRIELE INGLESE (\*)

**SOMMARIO.** - *In questo lavoro si considera il problema di ricostruire in  $[0, 1]$  una densità di probabilità  $u$  incognita di cui sono noti un numero finito di momenti e qualche informazione a priori di carattere locale (posizione e tipo di singolarità di  $u$  o  $\frac{du}{dx}$ ). Se l'informazione locale può essere riassunta a sua volta in una densità  $w$ , un modo naturale per stimare  $u$  consiste nel minimizzare una opportuna misura della discrepanza tra  $u$  e  $w$ . Abbiamo considerato a tale scopo l'entropia relativa e la distanza euclidea confrontando le corrispondenti soluzioni in una serie di esempi numerici.*

**SUMMARY.** - *The present paper deals with the reconstruction of an unknown probability density  $u$  in  $[0, 1]$  from a finite number of moments and some additional local a priori information (location and type of singularities of  $u$  or  $\frac{du}{dx}$ ). If the additional information may be represented by means of a density  $w$ , it is natural to select our estimator of  $u$  by minimizing some kind of discrepancy between  $u$  and  $w$  like euclidean distance or relative entropy. We compare the corresponding solutions in several numerical experiments.*

---

(\*) Indirizzi degli Autori: D. Fasino: Dipartimento di Matematica e Informatica, Università di Udine, via delle Scienze 206 33100 Udine (Italy), fax: 432-558499, e-mail: fasino@dimi.uniud.it. G. Inglese: CNR-Istituto di Analisi Globale e Applicazioni, Firenze, via S. Marta 13/a 50139 Firenze (Italy), fax: 55-475915, e-mail: gabriele@ds.iaga.fi.cnr.it

**Key words:** finite moment problems, generalized concepts of distance, maximum entropy.

## 1. Introduction

The *finite Hausdorff moment problem* (FHMP) consists in recovering an unknown probability density  $u$  in  $[0, 1]$  whose first  $m$  moments  $\int_0^1 x^k u(x) dx$ ,  $k = 1, \dots, m$ , are known to match (or to be ‘near’ in some sense to) a vector of given real numbers  $\mu \equiv (\mu_1, \dots, \mu_m)$ .

Let  $D^m \subset \mathbb{R}_+^m$  be the *convex hull* of the curve  $M = \{(t, \dots, t^m); t \in [0, 1]\}$  (i.e. the intersection of all the convex sets including  $M$ ).  $D^m$  is called in literature *m-moment space* and is known to be a convex *body* i.e. its interior ( $\text{int}D^m$ ) includes a ball of dimension  $m$  [1, thm. 7.3].

It is known (see [1]) that if  $\mu$  is outside  $D^m$ , the corresponding FHMP does not admit any solution, while if  $\mu \in \partial D^m$  ( $\partial D^m$  is the boundary of  $D^m$  and clearly has  $m$ -dimensional Lebesgue measure equal to zero) the only object having  $\mu_1, \dots, \mu_m$  as first moments is a (uniquely determined) convex combination of Dirac’s  $\delta$ .

The most interesting case is when  $\mu \in \text{int}D^m$ . In this case, the FHMP is a typical underdetermined inverse problem. Furthermore, as a consequence of the ill-posedness of the (infinite dimensional) Hausdorff moment problem (see [2]), any *accurate* computational approach to FHMP is severely ill-conditioned. Although we experienced that, whatever method we use, the error affecting the data is magnified approximatively by a factor  $e^{1.76 \dots m}$ , research about finite moment problem is not discouraged thanks to the following reasons:

1. smooth densities of physical interest have been observed ‘store the information’ mainly in their first moments [3];
2. the number of moments which allow to detect the shape of an unknown density can be reduced if other kind of information is available in addition.

Indeed, additional information (possibly taking the form of a prior density  $w$ ) is necessary when  $u$  is known to have singularities. Methods based on the minimization of some kind of generalized distance from  $w$  (see for example [4, 5]) seem to be quite natural and efficient in order to use both global (moments) and local (location and type of singularities of  $u$  and its derivatives) data. Here we concentrate our attention on the *minimum euclidean distance* (MED) and *minimum relative entropy* (MRE) methods, to be described in section 2. In section 3 we apply these methods to a moment problem arising in solid state physics. Section 4 deals with the error

propagation in MRE method.

## 2. MED and MRE solutions for moment problems

The *minimum euclidean distance* (MED) method for solving a FHMP with data  $\mu$  when a rough approximation (*prior density*)  $w \geq 0$  of the unknown is given, consists in minimizing the  $L^2$ -norm functional

$$N_w(u) = \int_0^1 (u - w)^2 dx \quad (2.1.a)$$

under the constraints

$$\int_0^1 x^k u(x) dx = \mu_k \quad k = 1, \dots, m. \quad (2.1.b)$$

It is easy to check that the minimizer (called *MED solution* of the FHMP or *estimator* of the unknown density) is unique and takes the form  $u_{Eucl}^{(m)} = w(x) + \sum_{k=0}^m a_k x^k$ . To the best of our knowledge this method has been used for the first time around forty years ago in the study of the thermodynamics of crystals [6].

The *maximum entropy* (ME) principle, introduced by Jaynes in 1957 [7], suggests a way to select one solution of an underdetermined inverse problem without improperly adding information. The principle of *minimum relative entropy* (MRE) (see [8]) is a generalization of the Jaynes principle and applies in cases when, in addition to the data, a *prior density*  $w > 0$  is given. In applying MRE to the FHMP we have to minimize the *relative entropy* functional (in literature also I-divergence, cross-entropy, directed divergence and so on)

$$E_w(u) = \int_0^1 u \ln \frac{u}{w} dx \quad (2.1.c)$$

subject to the constraints (2.1.b).

The existence of a unique, non-negative minimizer (MRE solution or estimator)  $u^{(m)}$ , when  $w$  is not too bad and  $\mu \in \text{int}D^m$ , is proved in [9]. If  $w(x) \equiv 1$  the minimizer is known as the maximum entropy solution of the corresponding FHMP (or ME estimator).

A formal usage of Lagrange multipliers rule furnishes the expression

$$u^{(m)}(x) = w(x) \exp\left(\sum_{k=0}^m a_k x^k\right).$$

The relation between the coefficients  $(a_0, \dots, a_m)$  and the moments  $(\mu_0, \mu)$  is known to be a diffeomorphism  $\Phi : \mathbf{R}^{m+1} \rightarrow \{(t, t\nu); \nu \in D^m; t \geq 0\}$  whose jacobian matrix is the Hankel matrix determined by the first  $2m$  moments of the MES (see [10]).

REMARK 2.1. MED is a linear method and so it is much simpler than MRE from a computational point of view. On the other hand, MED has some foreseeable drawback: the solution may assume negative values and the oscillations due to the Gibbs effect are more emphasized than in MRE. Clearly, MRE (MED) method is recommended if  $u$  is known to be the product (sum) of a singular part times (plus) a smooth one.

If the whole sequence of the moments of  $u$  is known and suitable further assumptions about  $u$  and  $w$  are given, the sequence  $u^{(m)}$  of MRE estimators converges to  $u$  for  $m \rightarrow \infty$ . To show this, we first observe that (2.1.c,b) can be rewritten as the problem of minimizing

$$\int_0^1 p \ln p \, dW \quad (2.2.a)$$

subject to the constraints

$$\int_0^1 x^k p(x) \, dW = \mu_k \quad k = 1, \dots, m \quad (2.2.b)$$

where  $p \equiv u/w$  and  $dW = w dx$  is the probability measure induced by the density  $w$ . If the function  $u$  has singularities but  $u/w$  is smooth enough, the procedure of estimating  $u$  by means of MRE solutions is supported by the following convergence result which is an immediate extension of the main theorem in [11]:

PROPOSITION 2.1. *Let  $w$  be a given positive density. If  $u$  is such that  $p \equiv u/w > 0$  is twice continuously differentiable in  $[0, 1]$ , then the sequence  $p^{(m)}$  of the minimizers of the problem (2.2) converges to  $p$  uniformly in  $[0, 1]$ .*

EXAMPLE 2.1. Consider the density  $u(x) = -c \ln|x - \frac{1}{3}| \exp(-20(x - \frac{1}{2})^2 + 5 \sin x)$  where  $c = .124657 \dots$ . Since the smooth part is very close to the exponential of a polynomial of degree 2 we are able to recover  $u$  very well from  $\mu_1, \mu_2$  if we know that there is a logarithmic

singularity in  $\frac{1}{3}$  (see figure 1.a). It is not surprising that the maximum entropy method (or any other which does not take account of additional information) is unable to detect the singularity and produces poor approximations. In figure 1.b we show the maximum entropy estimator of  $u$  computed by using 8 exact moments. In figures 1.c and 1.d we show MED reconstructions obtained by using 2 and 6 moments respectively: the quality is much lower than using MRE because of the multiplicative structure of  $u$ . In this context we say that our data are ‘exact’ when improvements in their accuracy leave the quality of the reconstruction unchanged.

REMARK 2.2. A convergence result, analogous to the one in proposition 2.1, is also true for MED estimators if  $u - w$  is bounded from below and suitably smooth. Nevertheless, it requires more information than in MRE case. In fact, if  $u = u_{smooth} + u_{sing}$  we have that  $u - w$  is smooth when  $w \equiv u_{sing}$ . On the other hand, if  $u = u_{smooth}/u_{sing}$  we get a smooth  $u/w$  from any  $w \propto u_{sing}$ .

REMARK 2.3. Neither the relative entropy  $E_w(u)$  (which is always non-negative and vanishes only for  $u = w$ ) nor its symmetrized form

$$S(w, u) \equiv E_w(u) + E_u(w)$$

is a metric [9]. Nevertheless such a quantity is often regarded (see for example [12, 13]) as a significant measure of the discrepancy between the densities  $w$  and  $u$ . Geometrical properties of the relative entropy in analogy with metrics are studied in [9, 4, 5]. A rigorous analysis of the numerical treatment of the optimization problem (2.1) is in [10] while preconditioning techniques to compute the MRE solution of a FHMP are introduced in [14].

REMARK 2.4. The relative entropy  $E_w(u)$  can be also used as a *stabilizing functional*. It means that the minimum of

$$E_w(u) + \lambda \left( \sum_{k=1}^m |\tilde{\mu}_k - \mu_k|^2 - \epsilon^2 \right)$$

is a regularized solution in the sense of Tikhonov. Such a point of view is deeply analyzed in [15].

### 3. A finite moment problem arising in solid state physics. Construction of a prior probability

The standard model of a (monoatomic) crystal consists in a three dimensional lattice where the equilibrium positions of the  $N$  atoms constituting the crystal are the lattice points. Physicists are interested in the determination of the *partition function*  $Q$  of the system because from the knowledge of  $\ln Q$  it is immediate to derive its thermodynamical functions (Helmoltz free energy, internal energy, heat capacity at constant volume, ...). We are about to see how the derivation of  $\ln Q$  reduces itself to the solution of a Hausdorff moment problem.

The potential energy is assumed (as in the Born-Von Karman model [16, ch. 5]) to have the form

$$\mathcal{V}(0) + \frac{1}{2} \sum_{j,k=1}^{3N} V_{jk} \xi_j \xi_k$$

where  $\xi_j$  are generalized coordinates. Observe that it is the Taylor expansion of a generic  $\mathcal{V}(\xi)$  near an equilibrium position, truncated at the first non-trivial term. It can be successfully used to investigate small vibrations of the crystal (low temperature case). The (hessian) matrix  $V$  depends on the physical properties of the crystal (distance between atoms, masses,...) but in any case it is assumed to be symmetric and positive definite. Hence small vibrations in the crystal can be exactly decomposed into  $3N$  *independent normal modes of vibration* with frequency  $\omega_\alpha$  ( $\alpha = 1, \dots, 3N$ ). So the system has been reduced to  $3N$  independent (quantal) one-dimensional harmonic oscillators whose energy levels are known to be  $E_{\alpha,n} = (n + \frac{1}{2})h\omega_\alpha$  ( $h$  is the Plank constant) with  $\alpha = 1, \dots, 3N$  and  $n$  non-negative integer. We can write down the partition function  $Q(N, T)$  of the system (see [16, (3.5)]) and get

$$Q(N, T) = \sum_{\alpha=1}^{3N} \sum_{n=0}^{\infty} e^{-E_{\alpha,n}\beta} = \sum_{\alpha=1}^{3N} e^{-\frac{h\omega_\alpha\beta}{2}} \frac{1}{1 - e^{-h\omega_\alpha\beta}}$$

where  $\beta = 1/(kT)$ ,  $k$  is the Boltzmann constant and  $T$  is the absolute temperature. It is known that the frequencies lie in an interval  $(0, \omega_{max})$  independent of  $N$ . Then we can take the thermodynamic

limit  $N \rightarrow \infty$  obtaining (the complete derivation of this formula can be found in [16, sect. 5.2]):

$$-\ln Q(T) = c_\infty \beta + \omega_{max} \int_0^1 (\ln(1 - e^{-h\beta\omega_{max}t}) + \frac{h\beta\omega_{max}t}{2}) \rho(t) dt \quad (3.1)$$

where  $\rho$  is the limit frequency distribution whose domain has been normalized to  $(0, 1)$ . The (probability) density  $\rho$  would completely determine the thermodynamics of the crystal but it is generally unknown and only the first few even moments can be obtained by algebraical operations on the *dynamical* matrix  $V$  in non-trivial cases. As a consequence, the theory that we have just sketched was considered quite formal and limited in its applications, until a practical procedure for computing a larger number of the even order moments of  $\rho$  was proposed [17] in the sixties. Such a procedure uses mainly the fact that under the condition  $n < N/T$  the spectral moments

$$\mu_{2n}(N) = \frac{1}{3N\omega_{max}^{2n}} \sum_{k=1}^{3N} \omega_k^{2n}$$

on a lattice made up of  $N \times N \times N$  cells are equal to those of  $\rho$  (i.e. after the themodynamical limit has been performed). The numbers  $\omega_k^2$  are simply the eigenvalues of the given matrix  $V^\dagger$  whose elements are  $V_{ij}^\dagger = V_{ij}/\sqrt{m_i m_j}$  where the positive number  $m_k$  indicates the mass of the  $k$ th oscillator. Hence, the computation of  $\mu_{2n}(N)$  consists in evaluating the trace of the matrix  $V^\dagger$ . The only bound on the reachable precision is given by the computer arithmetics.

In [18] the data obtained in [17] (with more than 20 reasonably exact digits), for a *face centered cubic* (fcc) crystal with nearest-neighbor interaction, are used to evaluate the integral in (3.1) by means of gaussian quadrature rules. Since only the even moments of  $\rho$  are available, the change of variable  $x = t^2$  is required, so that the unknown function is actually  $u(x) = \rho(\sqrt{x})/(2\sqrt{x})$ . The Christoffel numbers (abscissas  $x_j^{(n)}$  and weight  $p_j^{(n)}$  of the quadrature rule) of order  $n = 15$  are computed using 30 moments. Although the authors are not interested in inferring the shape of  $\rho$ , this is feasible by using their own data, thanks to the asymptotic properties of the Christoffel functions (see [19]) for the weight  $u$ . The dotted line in figure 2.a is made up by the points  $(x_j^{(15)}, y_j^{(15)})$  where

$y_j^{(15)} = 15p_j^{(15)} / (c_{15}\pi\sqrt{x_j^{(15)}(1-x_j^{(15)})})$ ,  $j = 1, \dots, 15$  where  $c_{15}$  is a normalization constant. It shows a good agreement with the ME estimator (solid line) obtained by using only 9 moments. It is remarkable that both agree with the qualitative behavior of  $u$  as derived in [20] by means of a different technique which does not involve moments.

We deal with an inverse problem whose largely incomplete data are partially completed by the *a priori* knowledge of the location of a number of singularities of the first derivative of the unknown function. In what follows we propose a way for using the information about the singularities in order to construct a prior probability and produce MRE and MED estimators by using the least possible number of moments. We stress the fact that using few moments turns out to be a good general strategy because of the just remarked numerical instability of the moment problems. The results of the numerical examples reported in this section are anyhow unaffected by noise. This is due to the fact that a *low* number of *sharp* data are handled with double precision arithmetics (on a DECstation 5000/33).

We have constructed a prior density  $w$  (plotted in figure 2.b) by using the additional information about the singularities of  $u'$  in the Van Hove critical points  $x_1 = \frac{1}{4}$ ,  $x_2 = \frac{1}{2}$  and  $x_3 = \frac{1}{2} + \frac{\sqrt{2}}{4}$ , see [21, 22, 23], and in the extremes of  $[0, 1]$  ( $u(0) = u(1) = 0$ ; see [16, ch. 5]):

$$w(x) = \begin{cases} \frac{1}{2} + \frac{\arcsin(8x-1)}{\pi}, & \text{if } x \in [0, x_1] \\ 1 + \sqrt{(x-x_1)(x_2-x)}, & \text{if } x \in [x_1, x_2] \\ 1 - \sqrt{(x-x_2)(x_3-x)}, & \text{if } x \in [x_2, x_3] \\ \frac{1}{2} + \frac{1}{\pi} \arcsin\left(\frac{2x-1-x_3}{1-x_3}\right), & \text{if } x \in [x_3, 1]. \end{cases}$$

In figures 3.a, 3.b and 3.c we compare  $wp^{(3)}$ ,  $wp^{(4)}$  and  $wp^{(5)}$  (dotted lines) with the same order ME estimators whose stabilization is observed starting from  $m = 5$ . Using MRE or MED method, stabilization begins for  $m = 3$ . In figure 3.d we show MRE (dotted) and MED (solid) solutions obtained for  $m = 5$ : they behave in a similar way so that in the present example MED is clearly a more convenient method.



Estimates of  $u$  based on Pollaczek polynomials and Christoffel functions are shown in [24] and [25] respectively.

In the *pre-computer era* Lax and Lebowitz [6] used MED method in the case of a two dimensional fcc crystal whose spectral density had been theoretically derived by Montroll (see [16]). They used  $m = 5$  moments and the prior knowledge about  $u$  consisted of the singularities in the Van Hove points and of the values  $u(0)$  and  $u(1)$ . From the same data, MRE estimators can be obtained which show a comparable quality.

REMARK 3.1. Suppose that the prior probability  $w$  is a convex combination of  $\delta$  distributions concentrating the mass in  $n < \frac{m}{2} + 1$  points in  $[0, 1]$ . In this case we can introduce a sequence of densities of the form  $e^{P_m^j(x)}$  weakly convergent to  $w$  for  $j \rightarrow \infty$ , where  $P_m^j(x)$  are polynomials of degree  $m$ . For each  $j \geq 0$  the MRE solution with prior probability  $e^{P_m^j(x)}$  takes the form  $e^{P_m^j(x)} e^{Q_m^j(x)} = e^{R_m^j(x)}$  where  $Q_m^j(x)$  and  $R_m^j(x)$  are polynomials of degree  $m$  too. Since for any  $j$  we have

$$\int_0^1 x^k e^{R_m^j(x)} dx = \mu_k \quad k = 1, \dots, m$$

it is easy to observe that  $R_m^j(x) = R_m(x)$  is independent of  $j$  and  $e^{R_m(x)}$  is simply the ME solution of the FHMP with data  $\mu$ . Hence, the additional information included in this kind of  $w$  cannot be used by applying the MRE principle.

#### 4. MRE solution from noisy data

Let  $\bar{\mu} \equiv (\bar{\mu}_1, \dots, \bar{\mu}_m)$  be our data vector and  $\mu \equiv (\mu_1, \dots, \mu_m)$  the (not available!) vector of the exact moments of the unknown function  $u$ . We suppose that  $\|\bar{\mu} - \mu\|_\infty < \epsilon$  and introduce the following convex, closed, non-empty set of probability densities

$$A_{\epsilon, \bar{\mu}} = \{u \text{ s.t. } \int_0^1 x^k u(x) dx \in [\bar{\mu}_k - \epsilon, \bar{\mu}_k + \epsilon] \quad k = 1, \dots, m\}.$$

The relative entropy functional  $E_w(u) = \int_0^1 u \log \frac{u}{w} dx$  is strictly convex (see for example [11, sect. 2]) and attains its unique minimum for  $u(x) = w(x)/e$ .

Hence, there is a unique minimizer  $\bar{u}(x) = w(x) \exp(\sum_{k=0}^m \bar{a}_k x^k)$  of the relative entropy functional in the set  $A_{\epsilon, \bar{\mu}}$ . In the sequel,  $\bar{u}$

will be called *MRE solution of the FHMP with noisy data  $\bar{\mu}$*  (or *noisy MRE solution*).

REMARK 4.1. Observe that the relative entropy assumes a very useful form in correspondence to the MRE solution  $v$  of a FHMP with data  $\nu$ :

$$f(\nu) \equiv E_w(v) = a_0(\nu) + \sum_{k=1}^m a_k(\nu)\nu_k.$$

REMARK 4.2. The set  $D^m \subset \mathbb{R}^m$  is known to get thinner and thinner for increasing  $m$ . More precisely, the width in the direction of the  $m$ -th axis is estimated by  $2^{-2m-2}$  [1, thm. 25.5]. Hence, we observe that:

1. the function  $f$  varies very fast in any direction;
2. the noisy data are probably out of  $D^m$ .

Now, suppose that the *a priori* bound  $E_w(u) \leq C$  is given. The convex set  $\{u : E_w(u) \leq C\}$  is known to be weakly compact in  $L^1$  (see [11, sect. 2]). Hence, the set  $B_C = \{\mu(u) : E_w(u) \leq C\}$  is a convex compact subset of the moment space  $D^m$ .

Let  $u^{(1)}$  and  $u^{(2)}$  be the MRE solutions corresponding to the moment vectors  $\mu^{(1)}$  and  $\mu^{(2)}$  both in  $B_{\epsilon, \bar{\mu}} = B_C \cap \{\|\mu(u) - \bar{\mu}\|_\infty \leq \epsilon\}$ . In order to discuss the stability of noisy MRE solutions, we evaluate the  $L^1$  size of the set  $B_{\epsilon, \bar{\mu}}$  by using the symmetrized cross-entropy. In fact, it is not difficult to check [26] that  $\|u - w\|_1^2 \leq 2S(w, u)$  as a consequence of the Jensen inequality. Moreover,

$$S(u^{(1)}, u^{(2)}) = \sum_{k=0}^m (a_k^{(2)} - a_k^{(1)})(\mu_k^{(2)} - \mu_k^{(1)}) \leq \|(a^{(2)} - a^{(1)})\|_2 \epsilon.$$

So we have

$$\|(a^{(2)} - a^{(1)})\|_2 \leq \|J\Phi(1, \bar{\mu})^{-1}\|_2 \epsilon$$

for a suitable  $\bar{\mu}$  belonging to the segment joining  $\mu^{(1)}$  and  $\mu^{(2)}$ . The term  $\|J\Phi(1, \bar{\mu})^{-1}\|_2$  takes the form (see [14])  $\text{constant} \cdot e^{3.25 \dots m}$  (such a constant depends on and is divergent with  $C$ ). Hence we have the usual bound (compare [27, sect. 4])

$$\|u^{(1)} - u^{(2)}\|_1 \leq \min\{2, \alpha(C)e^{1.76m}\epsilon\}. \quad (4.1)$$

The estimate above is marked by an unavoidable exponential term regardless of the explicit form of  $\alpha(C)$ . Hence, we do not try to get

it. The stability estimate (4.1) and the last remark about the width of  $D^m$  suggest that the noisy solution of a FHMP essentially matches the prior probability  $w$ . This is confirmed by numerical experiments.

We have performed many numerical tests in which the coefficients  $\bar{a}$  are computed by solving the nonlinear programming problem

$$\min \sum_{k=0}^m a_k \mu_k$$

subject to the constraints :

$$\int_0^1 w(x) e^{\sum_{j=0}^m a_j x^j} dx = 1$$

$$\mu_k \in [\bar{\mu}_k - \epsilon, \bar{\mu}_k + \epsilon] \quad \text{for } k = 1, \dots, m$$

$$\int_0^1 x^k w(x) e^{\sum_{j=0}^m a_j x^j} dx = \mu_k \quad \text{for } k = 1, \dots, m.$$

Indeed, as a rule, numerical methods for solving non-linear optimization problems perform better when nonlinearities are in the constraints and the objective function is as much linear as possible. So in our tests, we made use of the observation in Remark 4.1. In computations, IMSL routine DNCONF [28] with starting point  $a_0 = \ln 1$ ,  $\mu_0 = 1$ ,  $a_k = 0$  and  $\mu_k = \mu_k(w)$  for  $k = 1, \dots, m$  is used.

We report here the results obtained with two different sets of data.

TEST 1. We try to recover the smooth test function  $u(x) = c(2e^{-x^2} + \sin(3.5\pi x) + 1)$  (where  $c = .631076\dots$ ) from the knowledge of its first  $m = 8$  moments. The prior probability is chosen as  $w(x) \equiv 1$ . Figures 4a, 4b and 4c show the reconstructions with  $\epsilon = 10^{-5}, 10^{-4}, 1$  respectively. Observe that 8 moments affected by error with  $\epsilon = 10^{-5}$  give more or less the same maximum entropy estimator as 6 exact moments (see figure 4.d).

TEST 2. Let the test function be  $u(x) = .754154\dots w(x) \frac{(1+.5\sin(4\pi x))}{(1+x^2)}$  where  $w(x) = -\ln|x - \frac{1}{3}|$ . In figure 5a and 5b we show the MRE solution obtained from exact moments and error size  $\bar{\epsilon} = .3125 \cdot 10^{-3}$  respectively (with  $m = 6$ ). When the error size is greater than  $\bar{\epsilon}$ , our routine fails in getting the MRE solution: it means that the

data vector is out of  $D^m$  or too close to its boundary. If  $\epsilon > \bar{\epsilon}$  we can compute anyway the noisy MRE solution. Figures 5c (6 moments) and 5d (2 moments) seems to confirm that in this case the reconstruction is essentially  $w$ .

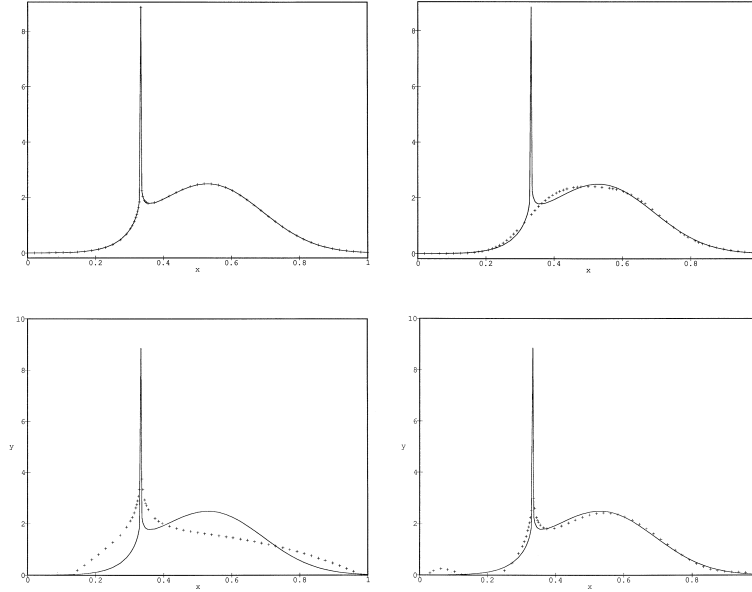


Figure 1: Reconstruction of  $u(x) = .124657 \dots w(x)e^{-20(x-\frac{1}{2})^2+5 \sin x}$  (with  $w(x) = -\ln|x - \frac{1}{3}|$ ) from exact data: (a) MRE,  $m = 2$ ; (b) ME,  $m = 8$ ; (c) MED,  $m = 2$ ; (d) MED,  $m = 6$ . The solid plot is the graph of  $u$ .

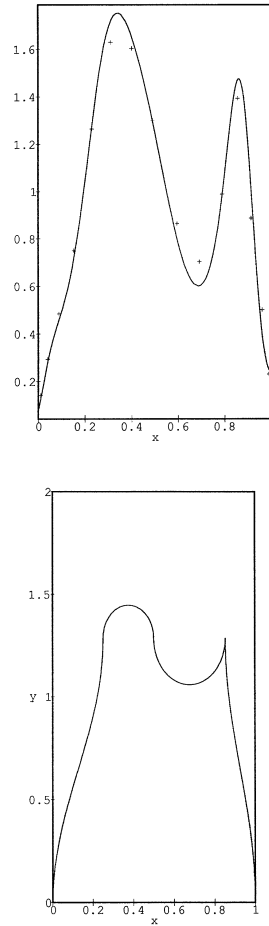


Figure 2: Reconstruction of the spectral density of a 3d fcc crystal from exact data: (a) ME estimator corresponding to  $m = 9$  (solid line) and Christoffel numbers approximation for  $m = 30$  (dots); (b) the proposed prior probability  $w$ .

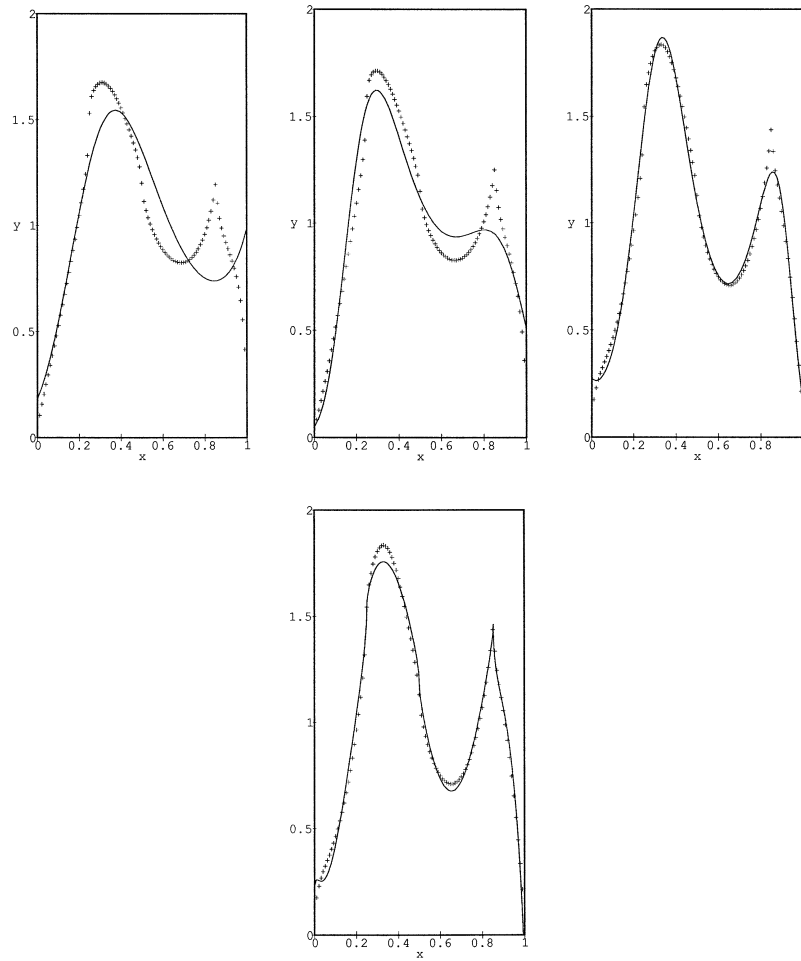


Figure 3: Reconstruction of the spectral density of a 3d fcc crystal from exact data: (a), (b) and (c) MRE (dotted) vs ME (solid) for  $m = 3, 4, 5$  respectively; (d) MRE (dotted) vs MED (solid) for  $m = 5$ .

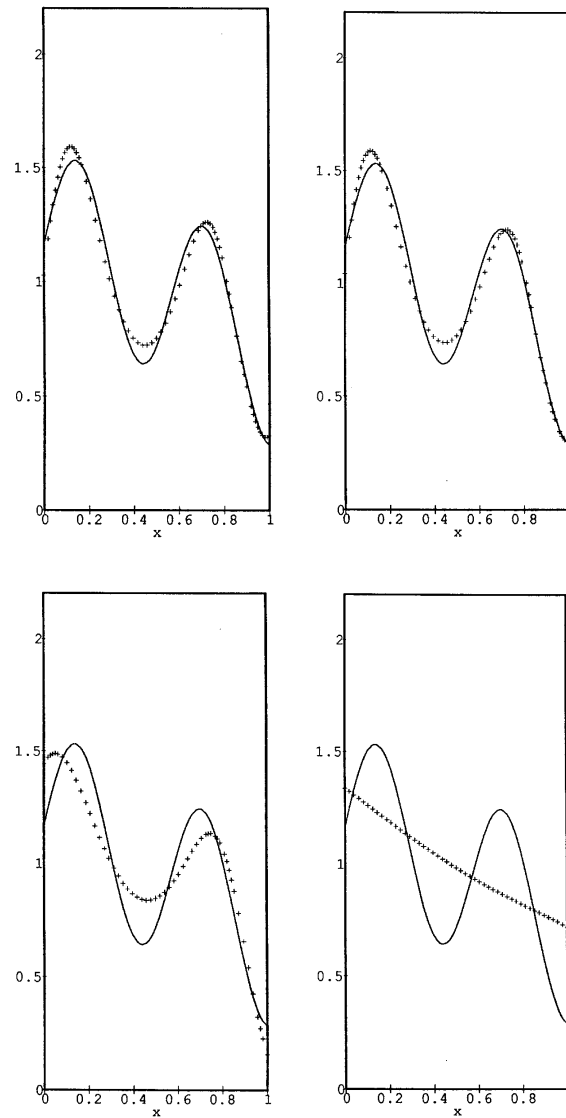


Figure 4: Reconstruction of  $u(x) = .631076 \dots (2e^{-x^2} + \sin(3.5\pi x) + 1)$  from noisy data and prior probability  $w(x) \equiv 1$ : (a) ME,  $m = 6$ , exact data; (b) noisy ME,  $m = 8$ ,  $\epsilon = 10^{-5}$ ; (c) noisy ME,  $m = 8$ ,  $\epsilon = 10^{-4}$ ; (d) noisy ME,  $m = 8$ ,  $\epsilon = 1$ . The solid plot is the graph of  $u$ .

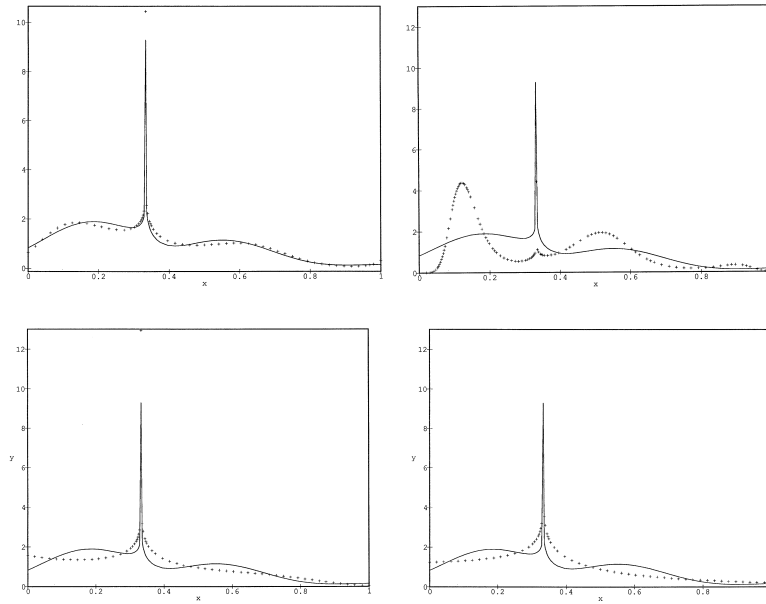


Figure 5: Reconstruction of  $u(x) = .754154\dots w(x) \frac{(1+.5\sin(4\pi x))}{1+x^2}$  (with  $w(x) = -\ln|x - \frac{1}{3}|$ ) from noisy data: (a) MRE,  $m = 6$ , exact data; (b) MRE,  $m = 6$ ,  $\bar{\epsilon} = .3125 \cdot 10^{-3}$ ; (c) MRE,  $m = 6$ ,  $\epsilon \geq \bar{\epsilon}$ ; (d) MRE,  $m = 2$ , same  $\epsilon$  as in (c).



## REFERENCES

- [1] KARLIN S. and SHAPLEY L. S., *Geometry of moment spaces*, Memoirs of the AMS **12** (1953) (AMS: Providence RI).
- [2] TALENTI G., *Recovering a function from a finite number of moments*, Inverse Problems, **3** (1987), 501-517.
- [3] FRENCH J. B., *Elementary principles of spectral distributions*, Theory and applications of moment problems in many-fermion systems, Dalton B J et al (eds) (New York: Plenum) (1979), 1-16.
- [4] JONES L. K. and TRUTZER V., *Computationally feasible high-resolution minimum distance procedures which extend the maximum-entropy method*, Inverse Problems **5** (1989), 749-766.
- [5] JONES L. K. and TRUTZER V., *On extending the orthogonality property of minimum norm solutions in Hilbert space to general methods for linear inverse problems*, Inverse Problems **6** (1990), 379-388.
- [6] LAX M. and LEBOWITZ J. L., *Moment Singularity Analysis of Vibration Spectra*, Phys. Rev. **96** (3) (1954), 594-598.
- [7] JAYNES E. T., *Information theory and statistical mechanics*, Physical Review **106** (4) (1957), 620-630.
- [8] JHONSON R. W. and SHORE J. E., *Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy*, IEEE Trans. Inf. Theory **IT-26** (1) (1980).
- [9] CSISZAR I., *I-divergence geometry of probability distributions and minimization problems*, The Annals of Probability, **1** (1) (1975), 146-158.
- [10] INGLESE G., *On the minimum relative entropy solutions of finite moment problems*, Numerical Functional Analysis and Optimization **16** (10) (1995), 1143-1153.
- [11] BORWEIN J. M. and LEWIS A. S., *Convergence of best entropy estimates*, SIAM J. Optimization **1** (2) (1991), 191-205.
- [12] FITZPATRICK B. G., *Bayesian analysis in inverse problems*, Inverse Problems **7** (1991), 675-702.
- [13] MACAULAY V. A. and BUCK B., *Linear inversion by the method of maximum entropy*, Inverse Problems **5** (1989), 859-873.
- [14] FASINO D., *Spectral properties of the Hankel matrices and numerical solutions of finite moment problems*, Journal of Computational and Applied Mathematics **65** (1995), 145-155.
- [15] AMATO U. and HUGUES W., *Maximum entropy regularization of Fredholm integral equations of the first kind*, Inverse Problems **7** (1991), 793-808.
- [16] HILL T. L., *An introduction to Statistical Thermodynamics*, Addison-Wesley: Reading MA, 1962.
- [17] ISENBERG C., *Moment calculations in lattice dynamics I (fcc lattice*

- with nearest-neighbor interactions*), Physical Review **132** (6) (1963), 2427-2433.
- [18] WHEELER J. C. and GORDON R. G., *Rigorous bounds for thermodynamic properties of harmonic solids*, The Journal of Chemical Physics **51** (12) (1969), 5566-5583.
- [19] LEIGHTON R. B., *The vibrational spectrum and specific heat of a fcc crystal*, Review of modern physics **20** (1) (1948), 165-174.
- [20] NEVAI P., *Orthogonal Polynomials*, Memoirs of the AMS **213** (1979) (AMS: Providence RI).
- [21] MEAD L. R. and PAPANICOLAU N., *Maximum entropy in the problem of moments*, J. Math. Phys. **25** (8) (1984), 2404-2417.
- [22] VAN HOVE L., *The occurrence of singularities in the elastic frequency distribution of a crystal*, Phys. Rev. **89** (6) (1953), 1189-1193.
- [23] ROSENSTOCK H. B., *Dynamics of Simple Lattices*, Phys. Rev. **97** (2) (1955), 290-303.
- [24] SCALAS E. and VIANO G. A., *The Hausdorff moments in statistical mechanics*, J. Math. Phys. **34** (12) (1993), 5781-5800.
- [25] INGLESE G., *Christoffel functions and finite moment problems*, Inverse Problems **11** (1995), 949-960.
- [26] CRAIN B. R., *An information theoretic approach to approximating a probability distribution*, SIAM J. Appl. Math. **32** (2) (1977), 339-346.
- [27] INGLESE G., *A note about the discretization of finite moment problems*, Inverse Problems **10** (1994), 401-414.
- [28] IMSL, *User's manual*, IMSL: Houston, 1987.

Pervenuto in Redazione il 20 Agosto 1996.