Société de Calcul Mathématique, S. A.
Algorithmes et Optimisation

Méthodes probabilistes pour l'analyse des incertitudes

liées à la sûreté des réacteurs nucléaires

L'Hypersurface Probabiliste

Nouvelle Construction

Rapport rédigé par

Olga Zeydina,

Ingénieur de Recherche, Société de Calcul Mathématique S. A.

en préparation de sa thèse de doctorat

"Méthodes probabilistes pour la Sûreté Nucléaire"

Thèse préparée à l'Université de Bretagne Sud,
Laboratoire de Mathématiques et Applications
Thèse codirigée par Emile Le Page et Bernard Beauzamy

Rapport no 2 adressé à

l'Institut de RadioProtection et de Sûreté Nucléaire

en application de la commande R50/11026029 du 29 novembre 2006

janvier 2007

Executive summary

We present here a new simplified and strengthened approach, in order to construct the Experimental Probabilistic Hypersurface (EPH), due to Beauzamy [BB1], 2004. This approach relies (as the previous one) upon the principle of maximum entropy, but the construction is made using discrete entropy, rather than continuous one. The main difference with the previous construction lies in the propagation of information, which is much simpler : it can be done in a very elementary way, so there is no more difficult tuning of the parameters.

In the one dimensional case, we give an explicit construction of the Hypersurface, in the case of 1, 2, 3, 4 measures, for the computational code communicated by IRSN.

In the many-dimensional case, we give all formulas for explicit computation of the EPH, at any point in the space.

Table of contents

I. Introduction

We give here a new construction of the Experimental Probabilistic Hypersurface, introduced by Bernard Beauzamy in [BB1] (2004). Recall that the EPH provides a way of "storing" probabilistic information : if an experiment, or a computational code, has been executed N times, it provides N results. But each of these results depends on the specific values given to each of the K parameters. The EPH gives a way of anticipating, with some probability, a result associated with values of the parameters which have never been met before.

In practice, the EPH consists in a collection of densities of probability, one above each point of the configuration space. The density is a Dirac measure (certainty) if the measure has been made at that point, otherwise, the density is less and less concentrated when we move further away from measure points.

The construction in [BB1] is based upon a general principle of maximal entropy (minimal information) : we should not add any unnecessary information when we build the EPH. We keep this fundamental principle here.

The key point in the construction is the propagation of information, from a measure point to any other point. In [BB1], this propagation depended upon two parameters, a and λ, and these parameters were tuned in comparison with the global characteristics of the EPH : see [BB1] and our first report [OZ1].

However, this computation of the parameters is quite delicate and complicated in practice : it depends on a prescribed distance between the empirical repartition function and the global empirical function, computed from all local densities. But the choice of a prescribed distance, or accuracy, is rather arbitrary and does not completely respect the principle of maximal entropy.

Here, instead, we introduce a new construction of EPH based upon a very simple principle for propagating information, from any measure point. As in [BB1], entropy increases linearly, but this increment depends only upon one parameter, and this parameter is tuned only with the rule of maximum entropy, with no computation of distance. The resulting construction is considerably simpler, both in theory and in computer implementation.

Also, our construction relies only upon discrete entropy, thus avoiding the delicate comparisons between continuous and discrete entropy which were needed in [BB1].

II. General construction

We give it first in the case of a one-dimensional code, that is depending on one parameter. A numerical example (built on the function indicated by IRSN) is given below. We start with theoretical considerations.

Let $\theta = g(x)$ be a one-dimensional computational code. We can think of θ as representing a temperature (see the work for Framatome-ANP, [BB1]), but it does not matter here.

Before any measure has been made, we know (this is expert knowledge) that this temperature, no matter what x is, must lie between two bounds $t_{\min}$ and $t_{\max}$. We know also that x may vary only in some bounded interval $\left[x_{\min}, x_{\max}\right]$: in practice, physically speaking, neither the parameter nor the result may take infinite values.

Before any measure is made, all we know is that the result will be in the interval $\left[t_{\min}, t_{\max}\right]$: this is represented by a uniform law on this interval. So, before any measure is made, at any point x, our guess for the result is a uniform law on the interval $\left[t_{\min}, t_{\max}\right]$.

Now, assume a measure has been made at a point $x = \xi_1$; it gives $\theta_1 = g(\xi_1)$. This corresponds to a new information. The EPH describes how this information propagates to any point $x \neq \xi_1$. The idea is that, if we move far from ξ_1, the information becomes less and less precise.

In Beauzamy's construction [BB1], information is characterized by entropy. The Minimal Information Lemma ([BB1]) shows that this entropy must increase linearly with the distance. But two definitions of entropy may be taken : discrete and continuous ; passing from one to the other is delicate.

Here, we will work only with discrete entropy : if a probability distribution (p_j), $p_j \geq 0$, $\sum_j p_j = 1$, is given, its entropy is defined as :

$$I = \sum_j p_j Log \frac{1}{p_j} \tag{1.1}$$

The minimal value of I is 0, when all p_j's are 0 except one, which is 1, and the maximal value, for a distribution over an interval, is obtained with the uniform distribution, when all p_j's are equal. So, if we set :

$$t_j = t_{\min} + \frac{j}{\nu}\left(t_{\max} - t_{\min}\right) \ , \ j = 0, \ldots, \nu \, , \tag{1.2}$$

the step of the division is :

$$\tau = \frac{t_{\max} - t_{\min}}{\nu} \, , \tag{1.3}$$

and the maximal entropy is :

$$I = Log(\nu) \, . \tag{1.4}$$

So, for a given point x, the entropy of the density function above this point is of the form

$$I(x) = \lambda d(x, \xi_1) \tag{1.5}$$

where λ is a constant to be determined (independent of x) and $d(x, \xi_1)$ denotes simply the distance between x and the measure point ξ_1, that is, in the one-dimensional case :

$$d(x,\xi_1) = |x - \xi_1|. \tag{1.6}$$

(Extensions to higher dimensions will be given below).

When the entropy $I(x)$ is fixed, the form of the probability distribution $\left(p_j(x)\right)$ follows. Indeed, Lemma 2 in [BB1] shows that, for given entropy, the probability with maximal variance is a "discrete Gaussian", that is of the form :

$$p_j(x) = \exp\left(-at_j^{\ 2} + b\right), \tag{1.7}$$

where a and b depend on x.

We first recall some items from the previous construction and we explain our modifications.

A. Previous construction

The continuous analogue, in [BB1], was Gaussian functions of the form :

$$f(t,x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left\{-\frac{(t-\theta_1)^2}{2\sigma^2}\right\},$$

with $\sigma = \lambda \cdot a^{d(x,\xi_1)} = \lambda \cdot a^{|x-\xi_1|}$, where $\lambda > 0$, $a > 1$.

In order to build the Hypersurface we had to find the best a and λ which satisfy the condition (see [BB1] and [OZ1]) :

$$\begin{cases} \left\|F_{emp}(t) - F_{EPH}(t)\right\|_{L_2} \le \varepsilon \\ Ent(EPH) \to \max \end{cases}$$

where :

$$F_{emp}(t) = \begin{cases} 0, & \text{if } t < \theta_1 \\ 1, & \text{if } t \ge \theta_1 \end{cases}$$

is the empirical repartition function, and :

$$F_{EPH}(t) = \frac{1}{x_{\max} - x_{\min}} \int_{x_{\min}}^{x_{\max}} \int_{t_{\min}}^{t} f(\tau,x)\, d\tau\, dx$$

is the repartition function deduced from all local densities.

The bound ε was arbitrarily fixed, with respect to the L^2 norm :

$$\varepsilon = \frac{1}{10}\left\|F_{emp}(t)\right\|_{L_2}$$

The entropy used was the continuous one :

$$Ent(EPH) = -\int f_{EPH}(t) \mathrm{Log}\, f_{EPH}(t)\ dt$$

In order to compute λ and a we tried successively different values with a trial-and-error method.

B. New construction

In the present construction, we will only use discrete functions. Instead of a continuous function, as a density, $f(t,x)$, we will use a collection of probabilities $p_j(x)$ (see formula (1.7)) :

$$p_j(x) = \exp\left(-at_j^{\ 2} + b\right),$$

where a and b depend on x.

This will allow us to work only with discrete entropy, thus avoiding the difficulties connected with the comparison between continuous and discrete entropy. This will also make the computations easier.

We can assume that, for all x, the discretization in t is always made the same way : it corresponds to the precision of the measure instrument. For instance, we can assume that, no matter what are the values of the input parameters, the output (as a temperature) is measured within 1°C.

This means that, for all x, the densities $p_j(x)$ will all be discretized the same way, as in (1.2) and (1.3) .

We assume first that the measure has been taken for $\xi_1 = 0$ and gave the value $\theta_1 = 0$. So the p_j are all symmetric : $p_{-j} = p_j$.

The local repartition functions $F(t,x)$ are replaced by the sequence of partial sums :

$$q_j(x) = \sum_{i=j_{\min}}^{j} p_i(x).$$

The global repartition function is therefore replaced by the partial sums :

$$G_j = \frac{1}{x_{\max} - x_{\min}} \int_{x_{\min}}^{x_{\max}} \sum_{i=j_{\min}}^{j} p_i(x) dx.$$

We have seen that the discrete entropy is given by the formulas (1.5) :

$$I(x) = \lambda d = \lambda \left|x - \xi_1\right|, \text{ with } \lambda > 0.$$

Note that λ depends on the discretization, namely on ν. We denote it by $\lambda(\nu)$. The choice of $\lambda(\nu)$ will be determined by the value of the entropy at the largest distance from ξ_1 to the endpoints, $x_{\min}$ and $x_{\max}$. We denote by $d_{\max}$ this largest distance. So we get the relation :

$$\lambda(\nu) = \frac{I_{\max}}{d_{\max}}.$$

The computation of the largest distance will be explained below.

But on the other hand the entropy is defined by formula (1.1) : $I = \sum_j p_j Log \frac{1}{p_j}$, so we can write :

$$\sum_j p_j \ Log \frac{1}{p_j} = \lambda(\nu) d. \tag{2.1}$$

But :

$$Log \frac{1}{p_j} = at_j^{\,2} - b,$$

and so (2.1) becomes :

$$a \sum_t t_j^{\,2} p_j - b = \lambda(\nu) d. \tag{2.2}$$

Since the p_j's are symmetric, $\sigma^2 = \sum_t t_j^{\,2} p_j$ (see [BB1]) , so we have :

$$a\sigma^2 + b = \lambda(\nu) d. \tag{2.3}$$

The condition $\sum_j p_j = 1$ gives :

$$e^b \sum_j e^{-at_j^2} = 1.$$

So :

$$b = Log \frac{1}{\sum_j e^{-at_j^2}}. \tag{2.4}$$

So (2.2) becomes :

$$a \frac{\sum_j t_j^{\,2} e^{-at_j^2}}{\sum_j e^{-at_j^2}} - Log \frac{1}{\sum_j e^{-at_j^2}} = \lambda(\nu) d. \tag{2.5}$$

When $\lambda(\nu)$ and d are known, this equation allows us to find a.

In practice, we can make some approximate computations, as follows. Recall that τ is the step of the subdivision in t (see (1.3)) :

$$\sum_j e^{-at_j^2} \approx \frac{1}{\tau}\int_{-\infty}^{+\infty} e^{-at^2}dt = \frac{1}{\tau}\cdot\sqrt{\frac{\pi}{a}} \tag{2.6}$$

$$\sum_j t_j^2 e^{-at_j^2} \approx \frac{1}{\tau}\int_{-\infty}^{+\infty} t^2 e^{-at^2}dt = \frac{\sqrt{\pi}}{2\tau a^{3/2}} \tag{2.7}$$

and (2.5) becomes :

$$\frac{1}{2} - Log\left(\tau\sqrt{\frac{a}{\pi}}\right) = \lambda(\nu)\,d\,,$$

that is :

$$Log\left(\tau\sqrt{\frac{a}{\pi}}\right) = \frac{1}{2} - \lambda(\nu)\,d\,,$$

which can be written :

$$\sqrt{\frac{a}{\pi}} = \frac{1}{\tau}\exp\left\{\frac{1}{2} - \lambda(\nu)\,d\right\},$$

and finally :

$$a = \frac{\pi}{\tau^2}\exp\{1 - 2\lambda(\nu)\,d\}\,, \tag{2.8}$$

and :

$$b = Log\left(\tau\sqrt{\frac{a}{\pi}}\right) = \frac{1}{2} - \lambda(\nu)\,d\,. \tag{2.9}$$

So now we have the explicit dependence of a and b as functions of $\lambda(\nu)$. We can rewrite the expression for $p_j(x)$ using formulas (2.8) and (2.9) :

$$p_j(x) = \exp\left(-\frac{\pi t_j^2}{\tau^2}\exp\{1 - 2\lambda(\nu)\,d\} + \frac{1}{2} - \lambda(\nu)\,d\right),\ j = 0,\ldots,\nu\,, \tag{2.10}$$

When $\xi_1 \neq 0$ and $\theta_1 \neq 0$ (general case) we obtain the formula :

$$p_j(x) = \exp\left(-\frac{\pi\left(t_j - \theta_1\right)^2}{\tau^2}\exp\left\{1 - 2\lambda(\nu)\left|x - \xi_1\right|\right\} + \frac{1}{2} - \lambda(\nu)\left|x - \xi_1\right|\right) \tag{2.11}$$

So, we obtain an expression for $p_j(x)$ which depends only from one parameter, namely $\lambda(\nu)$. It also depends on the discretization parameter τ, but this parameter is known.

C. Computation of λ in the case of 1 parameter : theoretical approach

1. Case of 1 measure

In order to compute the parameter $\lambda(\nu)$ we will use the fact that, before any measures were made, we had a uniform law between $t_{\min}$ and $t_{\max}$. It means that for all points x ($x \in [x_{\min}, x_{\max}]$) will have the same probability at t_j, namely $1/\nu$. And discrete entropy in this case will be equal (see [BB1]) :

$$I = \mathrm{Log}(\nu) = \mathrm{Log}\frac{t_{\max} - t_{\min}}{\tau}.$$

This is the local entropy in each point x (all of them are equal).

Since we do not have any measure, we have the worst information in each point, and the worst information gives, as we know, the biggest entropy. It means that for the situation when we make 1 or more measure at any point, we should not obtain an entropy bigger than $I = \mathrm{Log}(\nu)$.

When we make the first measure, the entropy changes at each point. At the measure point ξ_1 the entropy is equal to zero (see [BB1]) and it increases linearly when we are going away from ξ_1. And the biggest (maximum) entropy will be at the point which is the furthest from ξ_1.

But, as we said, the entropy at any point cannot be larger than $I = \mathrm{Log}(\nu)$, so we have the formula for $\lambda(\nu)$:

$$\mathrm{Log}(\nu) = \lambda(\nu) \cdot d_{\max}$$

that is :

$$\lambda(\nu) = \frac{\mathrm{Log}(\nu)}{d_{\max}} \tag{3.1}$$

where $d_{\max}$ is the biggest distance from ξ_1 to any point in the space, that is either $x_{\max}$ or $x_{\min}$.

We present this example with an illustration (see below), in which the observed value ξ_1 is closer to $x_{\min}$ than to $x_{\max}$. In this case $d_{\max} = |x_{\max} - \xi_1|$.

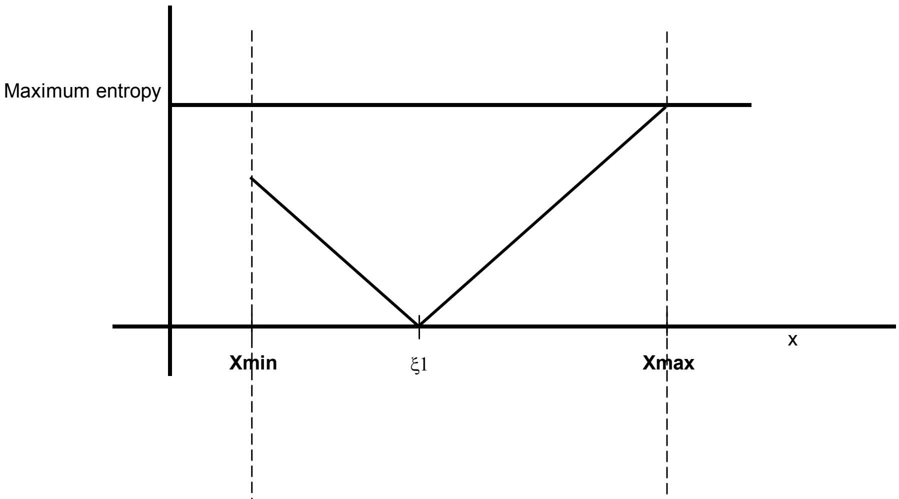


Figure 1 : increase of entropy with distance

The furthest distance $d_{\max}$ from ξ_1 to $x_{\max}$ (or $x_{\min}$) gives the best situation to compute λ. At this extreme point, the entropy should not be smaller than $I_{\max}$ (otherwise we add information).

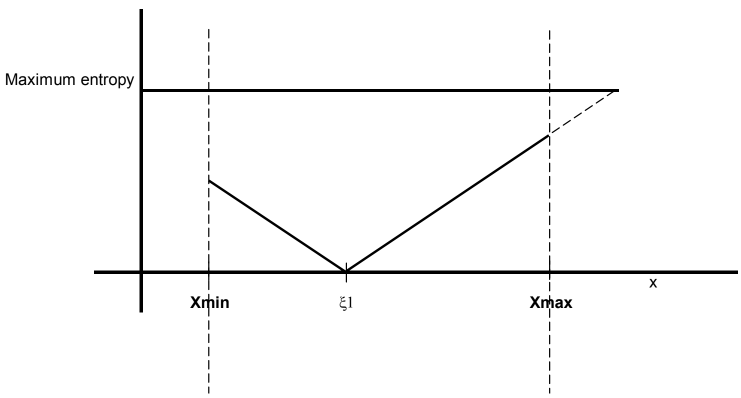


Figure 2 : the increase is too small

If the slope of increase is too large, then the entropy at the extreme point would be above the known value $\mathrm{Log}(\nu)$, which contradicts the fact that a new measure must add some information everywhere.

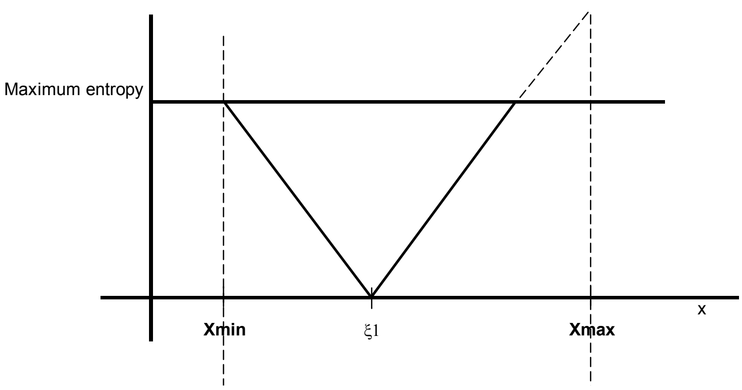


Figure 3 : the increase is too large

2. Case of 2 measures

We now study the case when we have two measures, at points ξ_1 and ξ_2. In order to compute $\lambda(\nu)$, we will use the same principle as before. We denote it by $\lambda_2(\nu)$. We now explain how to compute it.

First, we have to find which point, among ξ_1 and ξ_2, gives the furthest distance to x_{max} or x_{min}. Second, using this point we compute the entropy in the furthest point and obtain the entropy which gives the maximum, namely $\mathrm{Log}(\nu)$:

$$\mathrm{Log}(\nu) = \lambda_2(\nu) \cdot d_{\max} \tag{3.2}$$

where :

$$d_{\max} = \max\left\{\left|x_{\min} - \xi_1\right|, \left|x_{\max} - \xi_1\right|, \left|x_{\min} - \xi_2\right|, \left|x_{\max} - \xi_2\right|\right\}.$$

For example, in the picture below, the distance between ξ_2 and x_{max} is the biggest.
The entropy in the point x_{max}, with respect to ξ_2 will be the maximum, that is $\mathrm{Log}(\nu)$, so $\lambda_2(\nu) = \dfrac{\mathrm{Log}(\nu)}{\left|x_{\max} - \xi_2\right|}$.

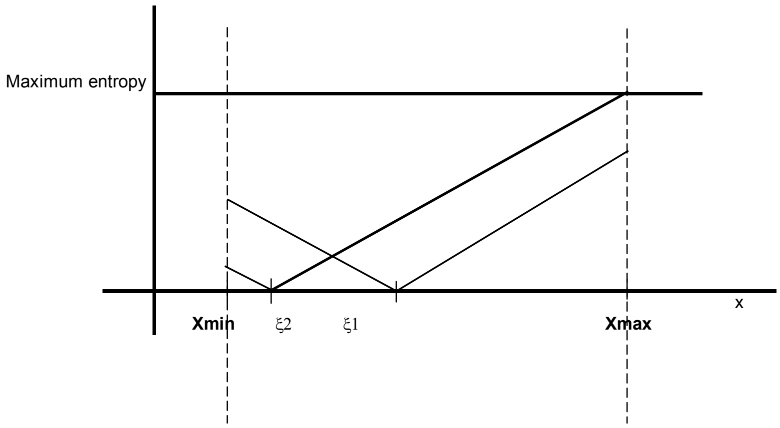


Figure 4 : increase of entropy with distance with two measures

Now, we have the representation of the resulting $p_j(x)$ for each $x \in \left[x_{min}, x_{max}\right]$ in the case of two measures. Recall that this density is given by a formula of the following type (see [BB1] and [OZ1]) :

$$p_j(x) = \frac{d_2}{d_1 + d_2} p_{1,j}(x) + \frac{d_1}{d_1 + d_2} p_{2,j}(x), \tag{3.3}$$

where $(p_{i,j})_j$ is the density corresponding to the $i-th$ measure, that is :

$$p_{i,j}(x) = \exp\left(-\frac{\pi\left(t_j - \theta_i\right)^2}{\tau^2} \exp\left\{1 - 2\lambda_2(\nu) d_i\right\} + \frac{1}{2} - \lambda_2(\nu) d_i\right) \tag{3.4}$$

and $d_i = d_i(x, \xi_i) = |x - \xi_i|$.

3. Case of N measures

Assume we made $N-1$ measures. All of them must give the same propagation of information and this propagation has to be independent of the order of the measures. In this case the formula for $\lambda_{N-1}(\nu)$ (where $\lambda_{N-1}(\nu)$ is for case with $N-1$ measure) is :

$$\lambda_{N-1}(\nu) = \frac{\mathrm{Log}(\nu)}{d_{\max}}$$

where $d_{\max} = \max\{|x_{\min} - \xi_n|, |x_{\max} - \xi_n|\}$ with $n = 1, \dots, N-1$. Denote it as $d_{\max, N-1}$

When we make a new measure ξ_N, we must have more information than before. Then, for N measures, $\lambda_N(\nu)$ has to become $\leq \lambda_{N-1}(\nu)$.

So, we compute the value : $d_{\max, N} = \max\{|x_{\min} - \xi_N|, |x_{\max} - \xi_N|\}$ and we compare $d_{\max, N-1}$ and $d_{\max, N}$.

If $d_{\max_N} > d_{\max_{N-1}}$ then we have to recompute $\lambda_N(\nu)$:

$$\lambda_N(\nu) = \frac{\mathrm{Log}(\nu)}{d_{\max_N}}.$$

If $d_{\max_N} \leq d_{\max_{N-1}}$ then we keep :

$$\lambda_N(\nu) = \lambda_{N-1}(\nu).$$

Now we have the representation of $p_j(x)$ for each $x \in [x_{\min}, x_{\max}]$ in the case of N measures :

$$p_j(x) = \frac{1}{\sum_{i=1}^{N} 1/d_i} \left(\frac{1}{d_1} p_{1,j} + \dots + \frac{1}{d_N} p_{N,j} \right), \tag{3.5}$$

where the $p_{i,j}$ are given by the formula (3.4).

This formula is simple, if one compares it with the first approach, given in [OZ1], in order to compute the unknown parameter λ.

Note also that here, we do not use any discretization in x, only a discretization in t.

D. Case of one parameter : Practical computation

1. Case of 1 measure

We take the example of a simple computational code suggested by IRSN :

$$g(x) = -1000\frac{\sin(x)}{(x+10)^2}$$

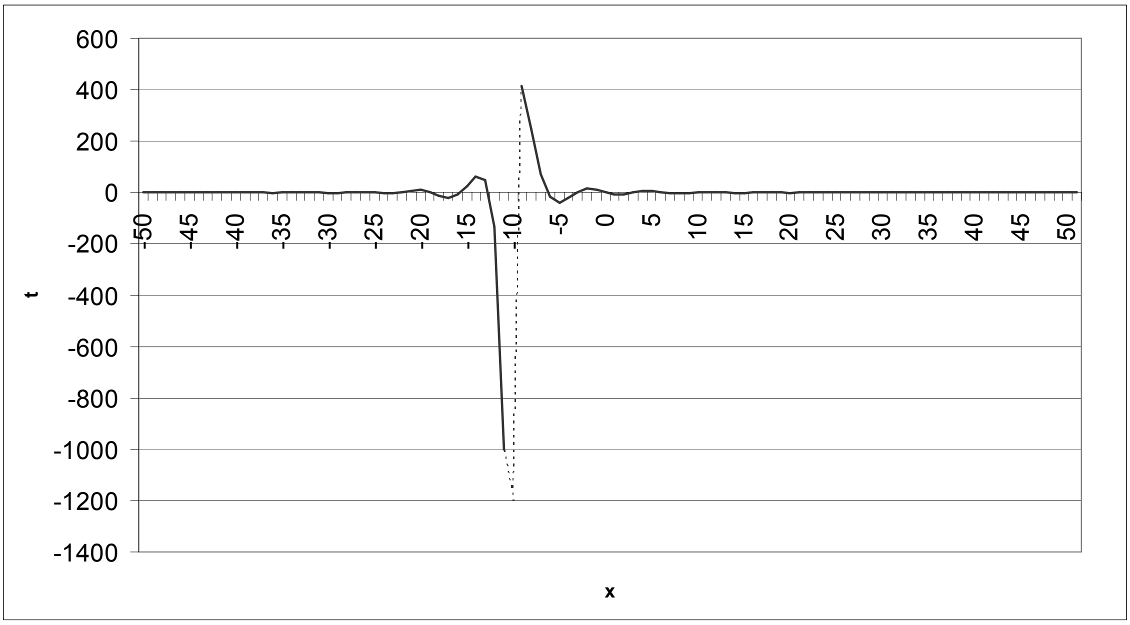
(one variable, one dimension)

We made only one measure:

$$\xi_1 = -4 \text{ and } \theta_1 = g(\xi_1) = -21\,.$$

First, we choose the boundaries for x and t. A computational code, or a physical experiment, admit only limited values for the input parameters and produce only limited values for the outputs.

The graph of this code has the form :



Graph 1 : the shape of the computational code

We take (arbitrarily) $x_{\min} = -50$, $x_{\max} = 50$ and $t_{\min} = -1100$, $t_{\max} = 500$. Discretization in t has step $\tau = 1$, then $\nu = \dfrac{t_{\max} - t_{\min}}{\tau} = 1600$.

In this case $d_{\max} = |x_{\max} - \xi_1| = 54$ and $\lambda_1(\nu) = \dfrac{\mathrm{Log}(\nu)}{d_{\max}} = 0.137$.

We put $\lambda_1(\nu)$ in our expression for $p_j(x)$ (formula (2.11)) :

$$p_j(x) = \exp\left(-\frac{\pi\left(t_j - \theta_1\right)^2}{\tau^2}\exp\left\{1 - 2\lambda_1(\nu)|x-\xi_1|\right\} + \frac{1}{2} - \lambda_1(\nu)|x-\xi_1|\right).$$

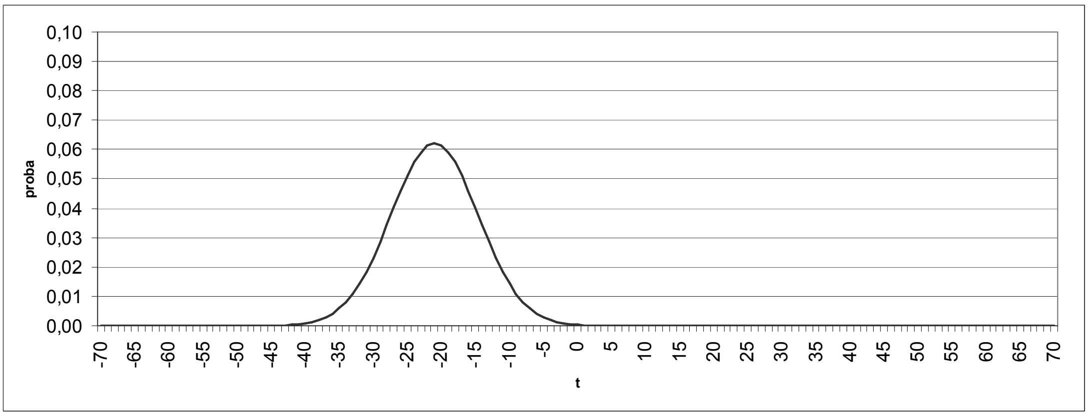
In this way we obtain the collection of densities above each point $x \in [x_{\min}, x_{\max}]$.
We have to take into account the fact we are working with bounded intervals,

$$t_j \in [t_{\min}, t_{\max}],\ x \in [x_{\min}, x_{\max}]$$

and we need to renormalize the formulas giving $p_j(x)$ so that the sum should be 1. Starting with the value given in (2.11), we introduce a normalization coefficient *c,* such that :

$$c \cdot \sum_{j=0}^{\nu} p_j(x) = 1 \quad , \quad c = \frac{1}{\sum_{j=0}^{\nu} p_j(x)} .$$

We present the graph of the normalized discrete function $p_j(x)$, for the point $x = 20$:



Graph 2 : the shape of the probability density for x=20

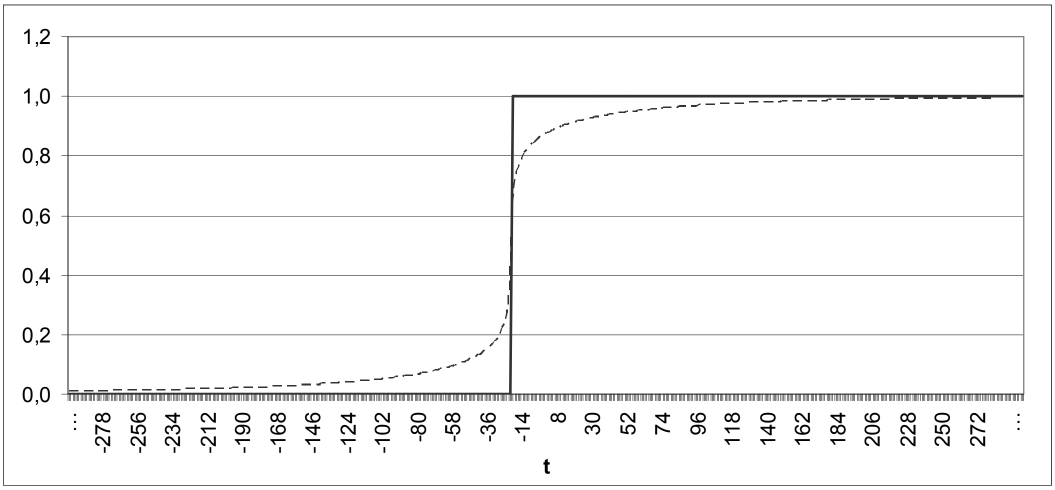
As an application, for example, we can compute the probability that at the point $x = 20$ our θ will be ≥ -10. This probability is given by the sum $\sum_{t_j=-10}^{500} p_j(x)$, and, with the numerical values taken here, we obtain :

$$\sum_{t_j=-10}^{500} p_j(x) = 0{,}051 .$$

The global repartition function G_j in the discrete approach has the form :

$$G_j = \frac{1}{x_{\max} - x_{\min}} \int_{x_{\min}}^{x_{\max}} \sum_{i=j_{\min}}^{j} p_i(x) dx$$

We build the graph of this function. Using an approximation to compute the integral, we obtain the graph :



Graph 3 : the shape of the repartition function

where G_j is a dotted line and empirical function $F_{emp}(t)$ is solid line.

In fact, here we do not need this function in order to compute λ (as we did in our continuous approach). So, this graph is just for presentation.

Also, we can draw a two-dimensional graph for the functions $p_j(x)$, when

$$t_j \in [t_{\min}, t_{\max}],\ x \in [x_{\min}, x_{\max}]$$

In this graph, a green line gives the density above the point $x_1 = -25$. The same density occurs above the point $x_1' = 17$, since it is symmetric of x_1, with respect to $x = -4$. The densities above points $x_2 = -17$ and $x_2' = 9$ are figured by red lines ; blue lines are for densities above points $x_3 = -11,\ x_3' = 3$.

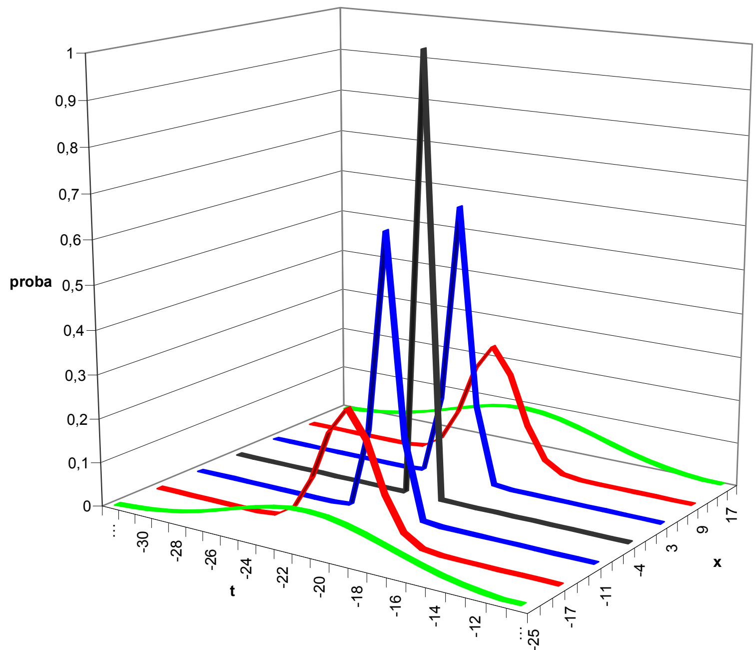


Figure 5 : the densities in the 2 dimensional space (in discrete points)

The continuous graph will be :

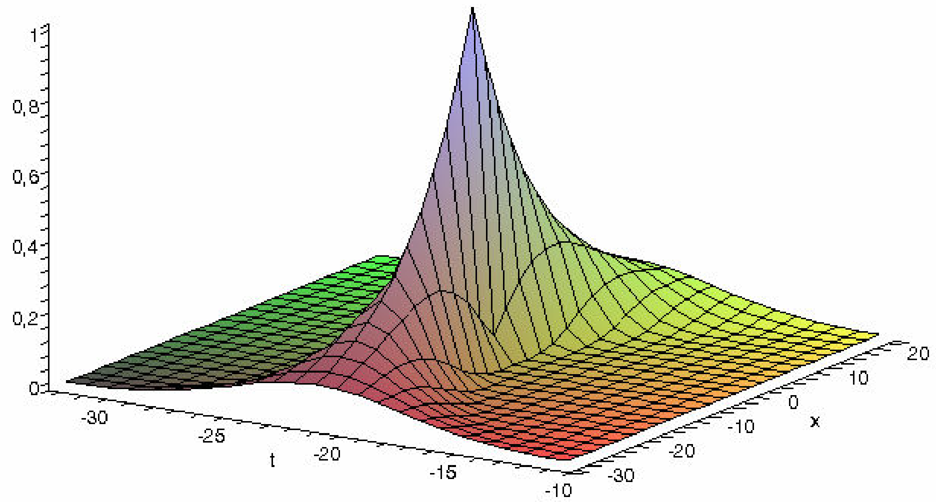


Figure 6 : the densities in the 2 dimensional space (continuous) ; one measure

4. Case of 4 measures

Let us take :

$$\xi_1 = -4\,,\ \xi_2 = -1\,,\ \xi_3 = 0\,,\ \xi_4 = 2$$

We obtain :

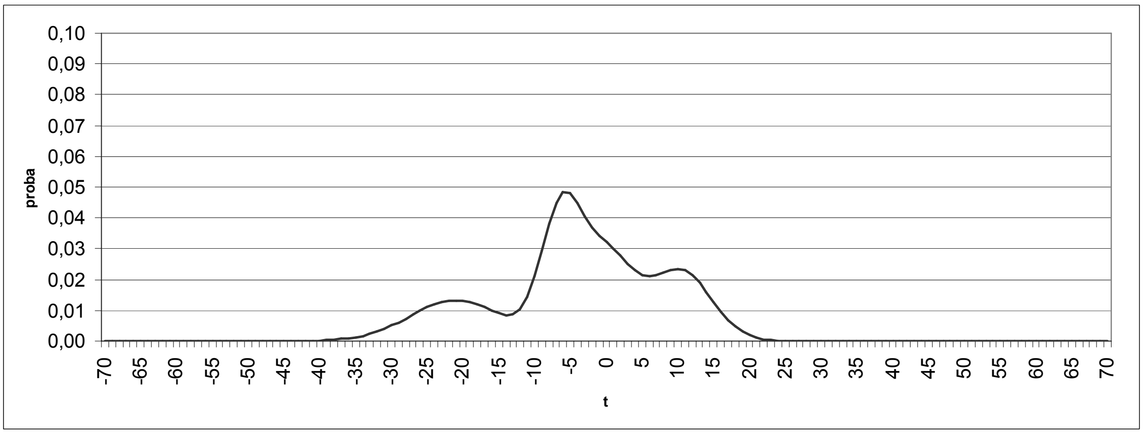
$$\theta_1 = g(\xi_1) = -21\,,\ \theta_2 = g(\xi_2) = 10.39\,,\ \theta_3 = g(\xi_3) = 0\,,\ \theta_4 = g(\xi_4) = -6.31$$

We take the same boundaries : $x_{\min} = -50$, $x_{\max} = 50$ and $t_{\min} = -1100$, $t_{\max} = 500$, discretization in t has step $\tau = 1$, $\nu = 1600$.

In the case of 4 measures the largest distance $d_{\max} = |x_{\max} - \xi_1| = 54$, we obtain :

$$\lambda_4(\nu) = \frac{\mathrm{Log}(\nu)}{d_{\max}} = 0.137$$

Let us compute the probability above the point $x = 20$. Using formula (3.5) with $N = 4$ we obtain the graph of function $p_j(x)$ for $x = 20$:



Graph 4 : the shape of the probability for x=20 with 4 measures

As an application, let us compute as before the probability that at the point $x = 20$ the temperature $\theta \geq -10$. This probability is :

$$\sum_{t_j=-10}^{500} p_j(x) = 0,777\,.$$

III. Building the EPH in the many-dimensional case

So far, we have worked on a one-dimensional space : the code depends only on one parameter. We will now see how the construction must be modified if the code depends upon many parameters. We start with two parameters.

We take the same notation as in [BB1] : the parameters are numbered with lower indexes and the measures with upper indexes.

A. Case of 2 parameters : theoretical construction

We now assume that our code depends on 2 parameters : $\theta = g(X) = g(x_1, x_2)$.

We made one measure at the point $A_1 = (\xi_1^{(1)}, \xi_2^{(1)})$ and we obtain $\theta_1 = g(\xi_1^{(1)}, \xi_2^{(1)})$.

So, we get one point A_1 in the space $\mathbb{R}^2$ and we build the Hypersurface in the space $\mathbb{R}^3$. In the general case, if we have K parameters, the equation $\theta = g(x_1 \dots x_K)$ determines a surface in the space $\mathbb{R}^{K+1}$ (see [BB1]).

In the case of 2 parameters (or, more generally, in the case of K parameters), each point of $\mathbb{R}^K$ has a density above it. These densities must be less and less concentrated, when we move further from the measure point : this is the same as in the case of one parameter. The entropy must increase linearly with the distance to the measure point, and the Gaussian distribution maximizes the variance for a given entropy : this is the same as in the one-parameter case.

So the densities are of the form :

$$p_j(x) = \exp\left(-a t_j^{\,2} + b\right),$$

as before.

The fact that the discrete entropy increases linearly with the distance gives the equation :

$$I(x_1, x_2) = \lambda(\nu)\, d(X, A_1) = \lambda(\nu) \sqrt{\left(x_1 - \xi_1^{(1)}\right)^2 + \left(x_2 - \xi_2^{(1)}\right)^2}$$

But on the other hand the entropy is defined by formula (1.1). So, we have :

$$\sum_j p_j \; Log \frac{1}{p_j} = \lambda(\nu) \sqrt{\left(x_1 - \xi_1^{(1)}\right)^2 + \left(x_2 - \xi_2^{(1)}\right)^2}$$

We make the same substitutions and approximate computations as we did for the case of 1 parameter and we obtain the expressions for a and b :

$$a = \frac{\pi}{\tau^2} \exp\left\{1 - 2\lambda(\nu) \cdot d(X, A_1)\right\},$$

and

$$b = \frac{1}{2} - \lambda(\nu) \cdot d(X, A_1)$$

Finally, we obtain the expression for the collection of densities for the case of 2 parameter and one measure :

$$p_j(x_1, x_2) = \exp\left(-\frac{\pi\left(t_j - \theta_1\right)^2}{\tau^2} \exp\left\{1 - 2\lambda(\nu)\, d(X, A_1)\right\} + \frac{1}{2} - \lambda(\nu)\, d(X, A_1)\right) \quad (4.1)$$

We define the boundaries : $x_1 \in \left[x_{1,\min}, x_{1,\max}\right],\ \ x_2 \in \left[x_{2,\min}, x_{2,\max}\right],\ \ \theta \in \left[t_{\min}, t_{\max}\right]$.

We compute $\lambda(\nu)$ using the same principle as before ; the furthest point gives the maximum entropy :

$$\mathrm{Log}(\nu) = \lambda(\nu) \cdot d_{\max},$$

where :

$$d_{\max} = \max\left\{\sqrt{\left(x_{1,\mathrm{m}} - \xi_1^{(1)}\right)^2 + \left(x_{2,\mathrm{m}} - \xi_2^{(1)}\right)^2}\right\},$$

and $x_{1,m}$ stands either for $x_{1,\max}$ or $x_{1,\min}$; the same for $x_{2,m}$

B. Case of K parameters and N measures

We now assume that our code depends on K parameters : $\theta = g(x_1, \ldots, x_K)$ and we made N measures. We want to build the Hypersurface at the point $X = \left(x_1, \ldots, x_K\right)$.

The first measure $A_1 = (\xi_1^{(1)} \ldots \xi_K^{(1)})$ gives us $\theta_1 = g(\xi_1^{(1)} \ldots \xi_K^{(1)})$, the second measure A_2 gives us $\theta_2 = g(A_2)$, and so on. In this way we obtain :

$$\begin{aligned} A_1 &= (\xi_1^{(1)} \ldots \xi_K^{(1)}) \to \theta_1 \\ A_2 &= (\xi_1^{(2)} \ldots \xi_K^{(2)}) \to \theta_2 \\ &\ldots \\ A_N &= (\xi_1^{(N)} \ldots \xi_K^{(N)}) \to \theta_N \end{aligned}$$

We have to define the bounds for each parameter at each point. We set :

$x_k \in \left[x_{k,\min}, x_{k,\max}\right]$ for $k = 1, \ldots, K$, and $\theta_1, \ldots, \theta_N \in \left[t_{\min}, t_{\max}\right]$.

The distance between points in the space with K parameters is the usual Euclidean distance. We set for $n = 1,\ldots,N$,

$$d_n = d(X, A_n) = \sqrt{\left(x_1 - \xi_1^{(n)}\right)^2 + \ldots + \left(x_K - \xi_K^{(n)}\right)^2}\,.$$

The general form of the collection of the densities with K parameters and N measures above each point X is :

$$p_j\left(X\right) = p_j\left(x_1, x_2, \ldots, x_K\right) = \gamma_1(X) \cdot p_{1,j}(X) + \ldots + \gamma_N(X) \cdot p_{N,j}(X) \qquad (5.1)$$

where $p_{n,j}(X)$ represents the contribution of the n -th measure, given by the formula :

$$p_{n,j}(X) = \exp\left(-\frac{\pi\left(t_j - \theta_n\right)^2}{\tau^2}\exp\left\{1 - 2\lambda_N\left(\nu\right)d_n\right\} + \frac{1}{2} - \lambda_N\left(\nu\right)d_n\right),\ n = 1,\ldots,N \qquad (5.2)$$

The coefficient of propagation $\lambda_N(\nu)$, depending on N (number of measures) and on ν (discretization) is given by the formula :

$$\lambda_N\left(\nu\right) = \frac{\mathrm{Log}\left(\nu\right)}{d_{\max}} \qquad (5.3)$$

where :

$$d_{\max} = \max\left\{\sqrt{\left(x_{1,\,\mathrm{m}} - \xi_1^{(n)}\right)^2 + \ldots + \left(x_{K,\,\mathrm{m}} - \xi_K^{(n)}\right)^2}\right\}\ ,\ n = 1,\ldots,N \qquad (5.4)$$

where $x_{k,m}$ stands either for $x_{k,\min}$ or for $x_{k,\max}$ and the maximum in (5.4) is taken over all possible combinations of max and min.

The coefficients $\gamma_n(X)$ in (5.1) must satisfy :

$$\gamma_1(X) + \ldots + \gamma_N(X) = 1. \qquad (5.5)$$

In a space with K parameters, each of them is of the form :

$$\gamma_n(X) = \frac{\frac{1}{d_n^K}}{\sum_{i=1}^{N}\frac{1}{d_i^K}}$$

(see [BB1], formula (10.2) and (10.3)).

So formula (5.1) takes a final form :

$$p_j(x_1, x_2, \ldots, x_K) = \frac{1}{\sum_{i=1}^{N} 1/d_i^K} \left(\frac{1}{d_1^K} p_{1,j} + \ldots + \frac{1}{d_N^K} p_{N,j} \right). \tag{5.6}$$

References

[BB1] The experimental probabilistic Hypersurface, Bernard Beauzamy, SCM SA, 2004.

[OZ1] L'Hypersurface Probabiliste. Rapport no 1, IRSN, SCM SA, 2006.