DETERMINATION OF A POWER DENSITY BY AN ENTROPY REGULARIZATION METHOD

OLIVIER PROT, MAÏTINE BERGOUNIOUX,
AND JEAN GABRIEL TROTIGNON

The determination of directional power density distribution of an electromagnetic wave from the electromagnetic field measurement can be expressed as an ill-posed inverse problem. We consider the resolution of this inverse problem via a maximum entropy regularization method. A finite-dimensional algorithm is derived from optimality conditions, and we prove its convergence. A variant of this algorithm is also studied. This second one leads to a solution which maximizes entropy in the probabilistic sense. Some numerical examples are given.

1. Introduction

In this paper, we present entropy regularization algorithms in order to determine electromagnetic wave propagation directions from the measurement of the six components of the electromagnetic field. Most of existing methods assume that the electromagnetic wave is similar to a single plane wave. In this case, for a fixed frequency, the electromagnetic field is fully described by the wave normal direction vector k. Nevertheless, this assumption is generally too restrictive. For a more realistic analysis of an electromagnetic wave in a plasma, Storey and Lefeuvre have introduced the concept of wave distribution function (WDF) [5]. This function describes the wave energy density distribution for every frequency and propagation modes. The WDF f is related to the spectral matrix V of the electromagnetic field component by a Fredholm integral of the first kind [5]

$$V(w) = \int q(k,w) f(k,w) dk, \qquad (1.1)$$

where k is the wave normal vector and w the frequency. In practice, the WDF is computed for the most significant frequency (often there is only one significant frequency) of the spectrogram. The integrating kernel q is a known function depending on the propagation media. The spectral matrix V is defined with the measured electromagnetic field $e : \mathbb{R}^+ \to \mathbb{R}^s$ by $V = \hat{e}(w)\hat{e}^*(w)$ where $\hat{e}$ denotes the Fourier transform of e and $s \in \mathbb{N}^*$ is the number of field components. It is a Hermitian matrix. This definition of the spectral matrix is an approximation of the real spectral matrix. Indeed, e is a random signal and

we should use mean values of e. Here we assume that e is deterministic. In what follows, we will identify the spectral matrix with a vector of $\mathbb{C}^n$, where $n = s^2$. Solutions of the inverse problem of determining f from measurements of V were proposed by Lefeuvre using a maximum entropy method [6]. However, for a fixed frequency, we have to solve the integral equation problem

$$V = \int q(x) f(x) dx, \tag{1.2}$$

where the unknown function f is nonnegative. This inverse problem is known to be ill-posed.

The concept of WDF can be transposed to the case of electromagnetic wave propagating in vacuum. Equation (1.1) remains valid if we use the vacuum kernels instead of the plasma kernels and if the electromagnetic wave has a single polarization mode [7]. The use of the WDF concept for electromagnetic wave propagating in vacuum has been studied for the interpretation of ground penetrating radar investigations, and, in particular, the one proposed for the NetLander mission to Mars [7]. The aim of this instrument was to explore the first kilometers of Mars subsurface to study its geological structure and look for water.

Here we use a maximum entropy regularization method to solve this problem. We minimize the quantity

$$\left\| V - \int q(x) f(x) dx \right\|_{\mathbb{C}^n}^2 + \mu H(f) \tag{1.3}$$

with the constraint $f \geq 0$, where H is a negentropic term (to be defined later) and μ a regularization parameter. In fact, H is not the negentropy in the probabilistic sense since f is not a density. But the minimization of H leads to a smooth solution. The main disadvantage of the maximum entropy solution of [6] is that the constraints on the solution are too strong. The regularization process provides a relaxed problem and the error we introduce allows to search a solution in a much wider domain. Moreover, we obtain a solution not far from the data that looks like a minimizer of H.

The maximum entropy regularization method is a useful tool to solve such inverse problems. Amato and Hughes [1] have studied the convergence of this method to show that it is a correct regularization process. This convergence is also studied in [3] by making a link with the classical Tikhonov method [11, 12]. A generalization is investigated in [8].

The mathematical model of the problem is described in Section 2. We define the negentropy in Section 3. In Section 4, we set the optimization problem: the feasible domain of this problem has to be relaxed to find optimality conditions. We present two algorithms, but the solutions we obtain do not minimize negentropy. So we modify the optimization problem in Section 5. Finally, we present numerical tests in Section 6.

2. Mathematical model

In this section, we present the mathematical model. We consider a measured space $(E, \mathscr{A}, \sigma)$, where E is a compact subset of $\mathbb{R}^p$, $p \geq 1$, and the measure σ verifies $\sigma(E) < +\infty$. The power density on $(E, \mathscr{A})$ can be defined as follows.

Definition 2.1. The power density is a couple (α, m), where $\alpha \in \mathbb{R}^+$ and m is a probability measure on $(E, \mathscr{A})$. Let $A \subset E$ be the power π_A of the subset A is given by

$$\pi_A = \alpha m(A). \tag{2.1}$$

The aim of this paper is to determine a power density (α, m) which verifies the equation

$$\alpha \int_E q \, dm = V, \tag{2.2}$$

where $V \in \mathbb{C}^n$ is known and $q \in L^2(E, \mathbb{C}^n, \sigma)$ is the integration kernel. In plasma physics, we have to solve this kind of problem to determine the power density distribution of an electromagnetic wave m and the total power α from the measurement of the electromagnetic field components. In this case, we typically have $n = 36$.

The set of probability measures is too large and we will only consider measures that are continuous with respect to the measure σ. We denote $\mathbb{H} = L^2(E, \mathbb{R}, \sigma)$ ($\subset L^1(E, \mathbb{R}, \sigma)$ thanks to our assumptions). For all $F \in \mathbb{H}$ and $F \geq 0$ σ-a.e., we can define a power density $(\alpha, (F/\alpha)d\sigma)$, where $\alpha = \|F\|_{L^1(E, \mathbb{R}, \sigma)}$ and $(F/\alpha)d\sigma$ is the measure of density F/α with respect to σ. Note that $F \mapsto \int_E \alpha q \, dm$ where $\alpha = \|F\|_{L^1(E, \mathbb{R}, \sigma)}$ and $m = F d\sigma/\alpha$ is a linear bounded operator from $\mathbb{H}$ to $\mathbb{C}^n$.

More generally, we consider a linear bounded operator $\psi = (\psi_1, \ldots, \psi_n) : \mathbb{H} \to \mathbb{C}^n$, $\psi^* : \mathbb{C}^n \to \mathbb{H}$ its adjoint operator and we assume that $R(\psi^*) \subset L^\infty(E, \mathbb{R}, \sigma)$ (R denotes the range). We have to solve

$$\psi[F] = V. \tag{2.3}$$

From the Riesz theorem, we deduced that there exist n functions $q_i \in L^2(E, \mathbb{C}, \sigma)$, $i = 1, \ldots, n$, such that $\psi_i[F] = \langle q_i, F \rangle_{\mathbb{H}}$. These functions are integration kernels, and we have

$$\forall l \in \mathbb{C}^n, \quad \psi^*(l) = \mathrm{Re}\left(\sum_{i=1}^n l_i \bar{q}_i \right), \tag{2.4}$$

where $\bar{q}_i$ denotes the conjugate complex of q_i. The condition $R(\psi^*) \subset L^\infty(E, \mathbb{R}, \sigma)$ is equivalent to

$$\forall i = 1, \ldots, n, \quad q_i \in L^\infty(E, \mathbb{C}, \sigma). \tag{2.5}$$

The problem of solving (2.3) is an ill-posed problem. Indeed, ψ is an operator from $\mathbb{H}$ (an infinite-dimensional Hilbert space) to the finite-dimensional Hilbert space $\mathbb{C}^n$. So, the operator ψ is not injective and there is no uniqueness of the solution (if it exists). In addition, we want to determine a solution which is also stable, that is continuous with respect to the data V. This will be crucial for the physical and numerical stability. To deal with this challenge, we use a maximum entropy regularization method.

The principle of Tikhonov's regularization method is to minimize the quantity

$$\|\psi[F] - V\|_{\mathbb{C}^n}^2 + \mu \Omega(F), \tag{2.6}$$

where μ is called the regularization parameter and Ω is a suitable regularizing functional. This method is equivalent to minimizing the functional Ω on the set $\{\|\psi[F] - V\|_{\mathbb{C}^n} \leq \delta(\mu)\}$ [1]. In this paper, we use an entropic term as the regularizing functional (see Section 3) and we restrict the domain to the Hilbert space $\mathbb{H}$ whereas usually, maximum entropy regularization is performed in L^1. We will see in the following section that there is no problem to define the entropy. First we recall some definitions.

3. About entropy

Let the function $\phi : \mathbb{R}^+ \to \mathbb{R}$ be defined by

$$\phi(x) = \begin{cases} x\ln(x) & \text{if } x > 0, \\ 0 & \text{else.} \end{cases} \tag{3.1}$$

The notation $m_1 \ll m_2$ means that the measure m_1 is absolutely continuous with respect to the measure m_2.

Definition 3.1. Let $f, g \in L^1(E, \mathbb{R}, \sigma)$ such that $f\, d\sigma$ and $g\, d\sigma$ are two probability measures and $f\, d\sigma \ll g\, d\sigma$. We define the relative information content of f with respect to g by

$$I(f, g) = \int_E f \ln \frac{f}{g}\, d\sigma. \tag{3.2}$$

If the condition $f\, d\sigma \ll g\, d\sigma$ is not verified, then $I(f, g) = +\infty$. If g is the noninformative probability density (see the definition below), then $I(f, g)$ is called the information content of f. The negentropy (negative entropy) of f is then defined by

$$H(f) = I(f, g). \tag{3.3}$$

The noninformative probability density is a known function of the model. Physically, it is the probability density of a noise measured in the system. For example, in the case of an electromagnetic wave in vacuum WDF, E is the unit sphere. Now, since isotropy occurs (there is no privileged direction of propagation for a plane wave in vacuum), we deduce that the noninformative density probability is constant over the unit sphere. So, we may assume that the noninformative probability density is given by

$$\forall x \in E, \quad g(x) = \frac{1}{\sigma(E)}. \tag{3.4}$$

For the sake of simplicity, we will suppose that $\sigma(E) = 1$, so the negentropy of f is

$$H(f) = \int_E \phi \circ f\, d\sigma. \tag{3.5}$$

Entropy of a probability density can be seen as a "distance" between f and the density of the noninformative probability g. To calculate the solution of the inverse problem, we minimize a functional involving negentropy. Thus, we determine the solution, which contains less information with respect to the density g. From the physical point of view,

this allows to preserve the physical significant information. So, it is primordial to know the noninformative density g quiet accurately. One can refer to [10] for the axiomatic derivation of the maximum entropy principle. The next lemma gives some properties of H.

LEMMA 3.2. *In the sequel, denote $K = \{f \in L^2(E, \mathbb{R}, \sigma) \mid f \geq 0 \ \sigma\text{-a.e.}\}$. For all $f \in K$, $-e^{-1} \leq H(f) < +\infty$. The functional $H : L^2(E, \mathbb{R}, \sigma) \to \mathbb{R}$ is lower semicontinuous (l.s.c.), strictly convex on K and verifies that for all $\varepsilon > 0$, for all $f \in K$ such that $f \geq \varepsilon \ \sigma\text{-a.e.}$, for all $g \in K$, for all $\lambda \in [0, 1]$,*

$$H(f + \lambda(g - f)) - H(f) = \lambda \int_E (1 + \ln f(x))(f - g)(x) d\sigma(x) + o(\lambda). \qquad (3.6)$$

Proof. Let $f \in L^2(E, \mathbb{R}, \sigma)$,

$$H(f) = \int_{\{x \in E \mid f(x) < 1\}} \phi \circ f(x) d\sigma(x) + \int_{\{x \in E \mid f(x) \geq 1\}} \phi \circ f(x) d\sigma(x); \qquad (3.7)$$

as

$$\begin{aligned} \forall x \in [0, 1], \quad -e^{-1} \leq x \ln x \leq 0, \\ \forall x \geq 1, \quad 0 \leq x \ln x < x^2, \end{aligned} \qquad (3.8)$$

so we deduce

$$H(f) \leq \|f\|_{L^2(E, \mathbb{R}, \sigma)} < +\infty, \qquad H(f) \geq -e^{-1}. \qquad (3.9)$$

The proof of the lower semicontinuity of functional H can be found in [1, 8].

We now prove (3.6): let $f \in K$ such that $f \geq \varepsilon$ let $g \in K$. We have for a.e. $x \in E$,

$$\phi(f(x) + \lambda(g(x) - f(x))) - \phi(f(x)) = \lambda(1 + \ln f(x))(g(x) - f(x)) + o(\lambda) \qquad (3.10)$$

since ϕ is derivable on $\mathbb{R}^{+*}$ and $\phi'(x) = 1 + \ln x$.

The functional H is strictly convex by the strict convexity of ϕ on $\mathbb{R}^+$. $\qquad \square$

4. A penalized problem

We now define the penalized cost functional or smoothing functional J_μ we want to minimize:

$$J_\mu : \mathbb{H} \longrightarrow \mathbb{R}, \qquad F \longmapsto \|V - \psi[F]\|_{\mathbb{C}^n}^2 + \mu H(F), \qquad (4.1)$$

where $\mu > 0$ is a regularization parameter. We have the following lemma.

LEMMA 4.1. *The functional J_μ is l.s.c. and strictly convex on K. In addition, if $\varepsilon > 0$, $f, g \in K$ with $f \geq \varepsilon \ \sigma\text{-a.e.}$, then*

$$\forall \lambda \in \,]0, 1[, \quad J_\mu(f + \lambda(g - f)) - J_\mu(f) = \lambda \langle DJ_\mu(f), g - f \rangle_{\mathbb{H}} + o(\lambda), \qquad (4.2)$$

where

$$DJ_\mu(f) = \mu(1 + \ln f) - 2\psi^*[V - \psi[f]], \tag{4.3}$$

$$J_\mu(g) \geq J_\mu(f) + \langle DJ_\mu(f), g - f \rangle_\mathbb{H}. \tag{4.4}$$

Proof. The functional $F \mapsto \|V - \psi[F]\|^2_{\mathbb{C}^n}$ is continuous on $\mathbb{H}$ by continuity of ψ, hence it is l.s.c. We conclude that J_μ is l.s.c. on K by Lemma 3.2. Furthermore, $F \mapsto \|V - \psi[F]\|^2_{\mathbb{C}^n}$ is Fréchet-differentiable and its gradient is $-2\psi^*[V - \psi[F]]$, so (4.2) is proved by Lemma 3.2 as well.

The functional J_μ is strictly convex on K by strict convexity of H and by convexity of the term $\|V - \psi[F]\|^2_{\mathbb{C}^n}$. So, we can write

$$J_\mu(f + \lambda(g - f)) < (1 - \lambda)J_\mu(f) + \lambda J_\mu(g),$$

$$\frac{J_\mu(f + \lambda(g - f)) - J_\mu(f)}{\lambda} < J_\mu(g) - J_\mu(f). \tag{4.5}$$

Equation (4.4) follows by taking the limit of the last equation when $\lambda \to 0$.   □

The penalized optimization problem stands as follows:
$(\mathscr{P}_\mu)$

$$\min J_\mu(F),$$

$$F \in K = \{f \in \mathbb{H} \mid f \geq 0 \ \sigma\text{-a.e.}\}. \tag{4.6}$$

The existence of a solution to problem $(\mathscr{P}_\mu)$ is not obvious since the cost functional is not coercive in $\mathbb{H}$. To illustrate this fact, we give a simple counterexample: we set $E = [0,1]$, σ is the Lebesgue measure on $[0,1]$ and $f : E \to \mathbb{R}$, $x \mapsto x^{-1/2}$. Then $H(f) = 2 < \infty$ and we can build a sequence $\{f_k\}_{k \in \mathbb{N}} \subset \mathbb{H}$ such that $f_k \to f$ a.e., $\|f_k\|_\mathbb{H} \to +\infty$, and $H(f_k) \to H(f) < +\infty$.

Nevertheless, since J_μ is strictly convex, the solution to $(\mathscr{P}_\mu)$ is unique if it exists. It is a function of $\mathbb{H} : F_{\mu,V}$. The power density can be obtained by setting $\alpha = \|F_{\mu,V}\|_{L^1(E,\mathbb{R},\sigma)}$ and $m = (F_{\mu,V}/\alpha)d\sigma$.

Therefore, we do not minimize the negentropy of Definition 3.1 since $F_{\mu,V}$ is not a probability density. Rigorously, the negentropy of the solution is $H(F_{\mu,V}/\alpha)$. Moreover, the cost functional does not verify (4.2) on the whole set K because ϕ is not derivable at 0. So, we have to modify this problem, taking a smaller feasible set. We study the modified problem in the next section to determine approximate first-order optimality conditions.

4.1. Relaxation of the feasible set. We just mentioned that $(\mathscr{P}_\mu)$ cannot be directly solved because J_μ is not coercive and (4.2) is not satisfied. So we choose a smaller feasible set to ensure (4.2). To deal with the lack of coercivity, we bound this new domain. For $0 < \varepsilon < T < +\infty$, we set

$$K_{\varepsilon,T} = \{f \in \mathbb{H} \mid \varepsilon \leq f \ \sigma\text{-a.e.}, \ \|f\|_\mathbb{H} \leq T\}. \tag{4.7}$$

It is a closed, convex subset of $\mathbb{H}$. The "relaxed" problem reads

$(\mathcal{P}_{\mu}^{\varepsilon,T})$

$$\min J_{\mu}(F), \qquad F \in K_{\varepsilon,T}. \tag{4.8}$$

THEOREM 4.2. *Problem (4.8) has a unique solution $F_{\mu}^{\varepsilon,T} \in K_{\varepsilon,T}$. A necessary and sufficient condition of optimality is*

$$\forall f \in K_{\varepsilon,T}, \quad \langle DJ_{\mu}(F_{\mu}^{\varepsilon,T}), f - F_{\mu}^{\varepsilon,T} \rangle_{\mathbb{H}} \geq 0, \tag{4.9}$$

where

$$DJ_{\mu}(F_{\mu}^{\varepsilon,T}) = -2\psi^*[V - \psi[F_{\mu}^{\varepsilon,T}]] + \mu(1 + \ln F_{\mu}^{\varepsilon,T}). \tag{4.10}$$

Proof. The existence and uniqueness of the solution is standard, see [2, Corollary III.20, page 46]. We call it $F_{\mu}^{\varepsilon,T}$. Then

$$\forall g \in K_{\varepsilon,T}, \ \forall \lambda \in [0,1], \quad J_{\mu}(F_{\mu}^{\varepsilon,T} + \lambda(g - F_{\mu}^{\varepsilon,T})) - J_{\mu}(F_{\mu}^{\varepsilon,T}) \geq 0. \tag{4.11}$$

With Lemmas 3.2 and 4.1, this is equivalent to

$$\forall g \in K_{\varepsilon,T}, \quad \langle DJ_{\mu}(F_{\mu}^{\varepsilon,T}), g - F_{\mu}^{\varepsilon,T} \rangle_{\mathbb{H}} \geq 0. \tag{4.12}$$

□

With the optimality condition (4.9), we may now construct the solution to $(\mathcal{P}_{\mu})$.

LEMMA 4.3. *If there exists $F_{\mu}^{\varepsilon,T} \in K_{\varepsilon,T}$ such that*

$$F_{\mu}^{\varepsilon,T} = \exp\left(-1 + \frac{2}{\mu}\psi^*[V - \psi[F_{\mu}^{\varepsilon,T}]]\right), \tag{4.13}$$

that is, $F_{\mu}^{\varepsilon,T}$ is a fixed point of a functional

$$\Gamma_{\mu} : \mathbb{H} \longrightarrow \mathbb{H}, \qquad F \longmapsto \exp\left(-1 + \frac{2}{\mu}\psi^*[V - \psi[F]]\right), \tag{4.14}$$

then for all $0 < \varepsilon' \leq \varepsilon$ and $T' \geq T$, $F_{\mu}^{\varepsilon,T}$ is the unique solution of $(\mathcal{P}_{\mu}^{\varepsilon',T'})$. Furthermore $F_{\mu}^{\varepsilon,T}$ is the unique solution of problem $(\mathcal{P}_{\mu})$ and we denote it by $F_{\mu,V}$.

Proof. Let $F_{\mu}^{\varepsilon,T} \in K_{\varepsilon,T}$ such that

$$F_{\mu}^{\varepsilon,T} = \exp\left(-1 + \frac{2}{\mu}\psi^*[V - \psi[F_{\mu}^{\varepsilon,T}]]\right) \ \sigma\text{-a.e.} \tag{4.15}$$

We get

$$\mu \ln(F_{\mu}^{\varepsilon,T}) = -\mu + 2\psi^*[V - \psi[F_{\mu}^{\varepsilon,T}]] \ \sigma\text{-a.e.} \tag{4.16}$$

So

$$\forall f \in K_{\varepsilon}, \quad \langle -2\psi^*[V - \psi[F_{\mu}^{\varepsilon,T}]] + \mu(1 + \ln(F_{\mu}^{\varepsilon,T})), f - F_{\mu}^{\varepsilon,T} \rangle_{\mathbb{H}} = 0, \tag{4.17}$$

and we see that $F_\mu^{\varepsilon,T}$ verifies (4.9). Therefore, $F_\mu^{\varepsilon,T}$ is the solution of (4.8). As $K_{\varepsilon,T} \subset K_{\varepsilon',T'}$ for all $0 < \varepsilon' \le \varepsilon$ and $T' \ge T$, we conclude that $F_\mu^{\varepsilon,T}$ is the solution of problem $(\mathcal{P}_\mu^{\varepsilon',T'})$. It is also the solution of problem $(\mathcal{P}_\mu)$. Suppose that $F' \in K$ exists such that $J_\mu(F') \le J\mu(F_\mu^{\varepsilon,T})$, then the function

$$\widetilde{F} = \frac{F_\mu^{\varepsilon,T} + F'}{2} \tag{4.18}$$

verifies $\widetilde{F} \ge \varepsilon/2 = \widetilde{\varepsilon}$ and $\|\widetilde{F}\|_\mathbb{H} \le (1/2)f(\|F_\mu^{\varepsilon,T}\|_\mathbb{H} + \|F'\|_\mathbb{H}) = \widetilde{T}$, so $\widetilde{F} \in K_{\widetilde{\varepsilon},\widetilde{T}}$. Since $F_\mu^{\varepsilon,T}$ is the solution of $(\mathcal{P}_\mu^{\widetilde{\varepsilon},\widetilde{T}})$, we have $J(F_\mu^{\varepsilon,T}) \le J(F')$ and we deduce that $F_\mu^{\varepsilon,T}$ is a solution of $(\mathcal{P}_\mu)$. Moreover, it is the unique solution of $(\mathcal{P}_\mu)$ since J_μ is strictly convex. $\square$

Lemma 4.3 also shows that if the functional Γ_μ has a fixed point, it is unique.

We are now able to find the solution as a fixed point. In the next subsection, we study the existence of a sequence that converges to this fixed point. That will be the essential tool to set an infinite-dimensional algorithm.

4.2. An infinite-dimensional algorithm. We define the sequence $\{F_k\}_{k \in N}$ of $\mathbb{H}$ by

$$F_0 \in \mathbb{H},$$
$$F_{k+1} = \exp\left(-1 + \frac{2}{\mu}\psi^*[V - \psi[F_k]]\right). \tag{4.19}$$

If it converges, the limit is a fixed point of the functional Γ_μ. It is also the solution to $(\mathcal{P}_\mu)$.

Lemma 4.4. *The functional Γ_μ is continuous from $\mathbb{H}$ to $L^\infty(E, \mathbb{R}, \sigma)$. Furthermore, the inequality*

$$\|\Gamma_\mu(F)\|_{L^\infty(E,\mathbb{R},\sigma)} \le \exp\left(-1 + \frac{2}{\mu}C_{\psi^*}[\|V\|_\mathbb{C}^n + C_\psi\|F\|_\mathbb{H}]\right) \tag{4.20}$$

is obtained.

Proof. We remark that $F \mapsto \exp(F)$ is continuous from $L^\infty(E, \mathbb{R}, \sigma)$ to itself and ψ^* is continuous from $\mathbb{C}^n$ to $L^\infty(E, \mathbb{R}, \sigma)$.

Inequalities are obtained with the continuity of operators ψ and ψ^*: there exist two constants C_ψ and C_{ψ^*} such that

$$\|\psi(F)\|_{\mathbb{C}^n} \le C_\psi\|F\|_\mathbb{H},$$
$$\|\psi^*(x)\|_{L^\infty(E,\mathbb{R},\sigma)} \le C_{\psi^*}\|x\|_{\mathbb{C}^n}. \tag{4.21}$$

Since the exponential function is nondecreasing, one obtains the results by injecting the last inequalities in the expression of Γ_μ. $\square$

Now we show the convergence of the sequence $\{F_k\}_{k \in N}$. The following lemma gives a condition on the regularization parameter μ which implies that the sequence $\{F_k\}_{k \in N}$ stays in a ball of fixed radius.

LEMMA 4.5. *Let F be such that* $\|F\|_{\mathbb{H}} \leq R$; *if*

$$\mu \geq \frac{2C_{\psi^*}[\|V\|_{\mathbb{C}^n} + C_\psi R]}{1 + \ln(R)}, \tag{4.22}$$

then $\|\Gamma_\mu(F)\|_{\mathbb{H}} \leq R$ *and* $\|\Gamma_\mu(F)\|_{L^\infty(E,\mathbb{R},\sigma)} \leq R$.

Proof. We have

$$\mu \geq \frac{2C_{\psi^*}[\|V\|_{\mathbb{C}^n} + C_\psi R]}{1 + \ln(R)},$$
$$\mu[1 + \ln(R)] \geq 2C_{\psi^*}[\|V\|_{\mathbb{C}^n} + C_\psi R], \tag{4.23}$$

so

$$R \geq \exp\left(-1 + \frac{2}{\mu}C_{\psi^*}[\|V\|_{\mathbb{C}^n} + C_\psi R]\right), \tag{4.24}$$

and we deduce the two inequalities.   □

We will use a fixed point criterion: if the functional Γ_μ is a contraction, then the sequence is converging. In the following lemma, we give a condition on μ for the sequence to converge.

LEMMA 4.6. Γ_μ *is Fréchet differentiable on* $\mathbb{H}$ *and its derivative is*

$$d\Gamma_\mu(F) \cdot h = -\frac{2}{\mu}\Gamma_\mu(F)\psi^* \circ \psi[h]. \tag{4.25}$$

Let $\|F_0\|_{\mathbb{H}} \leq R$. *If* μ *verifies (4.22) and if*

$$\mu > 2R^2 C_{\psi^*} C_\psi, \tag{4.26}$$

then the sequence $(F_k)_{k \in \mathbb{N}}$ *converges in* $\mathbb{H}$ *and in* $L^\infty(E,\mathbb{R},\sigma)$ *to the unique fixed point* $F_{\mu,V}$ *of* Γ_μ.

Proof. The functional Γ_μ is differentiable since $F \mapsto \exp(F)$ is differentiable from $L^\infty(E,\mathbb{R}, \sigma)$ to itself, and $F \mapsto \psi^*[\psi[F]]$ is linear continuous from $\mathbb{H}$ to $L^\infty(E,\mathbb{R},\sigma)$. By Lemma 4.5, we deduce that for all $k \in \mathbb{N}$, $0 \leq F_k \leq R$. Furthermore, (4.26) leads to

$$\sup_{\{F \in \mathbb{H} \mid 0 \leq F \leq R\}} \|d\Gamma_\mu(F)\|_{\mathscr{L}(L^\infty, L^\infty)} < 1. \tag{4.27}$$

We conclude using the Banach fixed point theorem on the complete set $\{F \in \mathbb{H} \mid 0 \leq F \leq R\}$ with the distance induced by L^∞. The sequence converges to the unique fixed point $F_{\mu,V} \in \{F \in \mathbb{H} \mid 0 \leq F \leq R\}$ of Γ_μ strongly in $L^\infty(E,\mathbb{R},\sigma)$ and in $\mathbb{H}$ (by compactness).   □

We may summarize in the following theorem.

THEOREM 4.7. *If μ verifies (4.22) and (4.26) (i.e., μ large enough), then problem ($\mathscr{P}_\mu$) has a unique solution $F_{\mu,V}$ limit of the sequence $\{F_k\}_{k\in\mathbb{N}}$ defined by*

$$F_0 \in \{F \in \mathbb{H} \mid 0 \le F \le R\},$$

$$F_{k+1} = \exp\left(-1 + \frac{2}{\mu}\psi^*[V - \psi[F_k]]\right) = \Gamma_\mu(F_k). \tag{4.28}$$

The convergence stands in L^∞ and in $\mathbb{H}$.

Shortly speaking, we have an infinite-dimensional algorithm that converges to the solution of the maximum entropy regularization problem for a fixed parameter μ great enough. Theorem 4.7 shows that the solution $F_{\mu,V}$ of ($\mathscr{P}_\mu$) is obtained as the limit of the sequence $\{F_k\}_{k\in N}$ and belongs to $L^\infty(E, \mathbb{R}, \sigma)$.

However this algorithm is not useful from the numerical point of view. Indeed, it is an infinite-dimensional one and an inappropriate discretization process may lead to slow computations. Nevertheless, we are able to derive a finite-dimensional algorithm from the optimality condition (4.9). This is the aim of the next subsection.

4.3. A finite-dimensional algorithm. Lemma 4.3 suggests to look for the solution $F_{\mu,V}$ of problem ($\mathscr{P}_\mu$) as $F_{\mu,V} = G_{\mu,V}$, where

$$G_{\mu,V} = \exp\left(-1 + \frac{2}{\mu}\psi^*[\lambda]\right), \tag{4.29}$$

where $\lambda \in \mathbb{C}^n$ has to be determined. The next lemma gives a sufficient optimality condition on λ to solve ($\mathscr{P}_\mu$). In addition, we have an analytic expression for the solution.

LEMMA 4.8. *Let $\lambda \in \mathbb{C}^n$ such that*

$$\lambda = V - \int_E q(\sigma)\exp\left(-1 + \frac{2}{\mu}\psi^*[\lambda](\sigma)\right)d\sigma, \tag{4.30}$$

then the function $G_{\mu,V} \in \mathbb{H}$, defined by (4.29), is the unique solution of ($\mathscr{P}_\mu$).

Proof. By the definition of $G_{\mu,V}$ and with the assumption $R(\psi^*) \subset L^\infty(E, \mathbb{R}, \sigma)$, there exist $\varepsilon > 0$ and $T > \varepsilon$ such that $G_{\mu,V} \in K_{\varepsilon,T}$. Writing the expression of $F_{\mu,V}$ in (4.9), we can see that it is verified. So $F_{\mu,V}$ is the unique solution of problem (4.8). With Lemma 4.3 and the strict convexity of J_μ, it follows that it is the unique solution of ($\mathscr{P}_\mu$). ☐

Therefore, we only need to find the value of $\lambda \in \mathbb{C}^n$ to determine $G_{\mu,V}$. So the problem turns to be a *finite-dimensional* one. We define the sequence $\{\lambda_k\}_{k\in\mathbb{N}}$ of $\mathbb{C}^n$ as

$$\lambda_0 \in \mathbb{C}^n,$$

$$\lambda_{k+1} = \gamma_\mu(\lambda_k) \stackrel{\text{def}}{=} V - \int_E q(\sigma)\exp\left(-1 + \frac{2}{\mu}\psi^*[\lambda_k](\sigma)\right)d\sigma. \tag{4.31}$$

The function $\gamma_\mu : \mathbb{C}^n \to \mathbb{C}^n$ defined in (4.31) is differentiable and its derivative is

$$d\gamma_\mu(\lambda) \cdot h = -\frac{2}{\mu}\int_E q(\sigma)\bar{q}^t(\sigma)\exp\left(-1 + \frac{2}{\mu}\psi^*[\lambda](\sigma)\right) \cdot h \, d\sigma; \tag{4.32}$$

thus

$$||d\gamma_\mu(\lambda_k) \cdot h||_{\mathbb{C}^n} \le \frac{2}{\mu} \rho_\mu^{(k)} ||h||_{\mathbb{C}^n}, \qquad (4.33)$$

where $\rho_\mu^{(k)}$ is the spectral radius of matrix $M_\mu^{(k)}$ of dimension (n,n) such that for all $1 \le i,j \le n$ (we use the Euclidean norm on $\mathbb{C}^n$),

$$M_{\mu,ij}^{(k)} = \int_E q_i(\sigma)\bar{q}_j(\sigma) \exp\left(-1 + \frac{2}{\mu}\psi^*[\lambda_k](\sigma)\right) d\sigma. \qquad (4.34)$$

Using the Frobenius norm of matrix $M_\mu^{(k)}$, we have the inequality

$$\rho_\mu^{(k)} \le ||M_\mu^{(k)}||_{Fr} \le \exp\left(-1 + \frac{2}{\mu} C_{\psi^*} ||\lambda_k||_{\mathbb{C}^n}\right) \sqrt{\sum_{1 \le i,j \le n} \left| \int_E q_i(\sigma)\bar{q}_j(\sigma) d\sigma \right|^2}. \qquad (4.35)$$

The sequence $\{\lambda_k\}_{k \in \mathbb{N}}$ converges to the fixed point of γ_μ only if $\rho_\mu^{(k)}$ is small enough for any k. So we cannot use it to calculate λ. However, we are able to construct another sequence converging to λ noting that for all $\tau > 0$, $\lambda = \gamma_\mu(\lambda)$ is equivalent to $\lambda = \lambda - \tau[\lambda - \gamma_\mu(\lambda)]$. So we can obtain λ as limit of the sequence $\{l_k\}_{k \in \mathbb{N}}$

$$
\begin{aligned}
l_0 &\in \mathbb{C}^n, \\
l_{k+1} &= l_k - \tau[l_k - \gamma_\mu(l_k)].
\end{aligned}
\qquad (4.36)
$$

This sequence will be used to determinate the solution practically. If τ is chosen small enough and μ is great enough, it converges.

LEMMA 4.9. *Assume that for every $k \in \mathbb{N}$, the spectral radius of $M_\mu^{(k)}$ is less than $m_\mu \in \mathbb{R}^{+*}$; then the sequence l_k converges if $0 < \tau < 1/(1 + (2/\mu)m_\mu)$.*

Proof. The matrix $M_\mu^{(k)}$ is Hermitian of nonnegative type since for all $\sigma \in E$, $q(\sigma)\bar{q}^t(\sigma)$ is nonnegative Hermitian and $\exp(-1 + (2/\mu)\psi^*[\lambda_k](\sigma)) > 0$. Let $h_\tau : \mathbb{C}^n \to \mathbb{C}^n$ be the function $\lambda \mapsto \lambda - \tau[\lambda - \gamma_\mu(\lambda)]$. The derivative of h_τ is

$$\forall \lambda, v \in \mathbb{C}^n, \quad dh_\tau(\lambda) \cdot v = \left([1-\tau]I - \frac{2}{\mu}\tau M_\mu\right)v. \qquad (4.37)$$

Since $M_\mu^{(k)}$ is Hermitian, there exists an orthogonal basis of eigenvectors and we call $B^{(k)}$ the transition matrix. So $B_*^{(k)}([1-\tau]I - (2/\mu)\tau M_\mu^{(k)})B^{(k)} = [1-\tau]I - (2/\mu)\tau\Delta_\mu^{(k)}$, where $(2/\mu)\tau\Delta_\mu^{(k)}$ is a diagonal matrix with positive elements. Thus, the spectral radius of $[1-\tau]I - (2/\mu)\tau M_\mu^{(k)}$ is strictly less than $1 - \tau$ since the spectral radius of $M_\mu^{(k)}$ is less than m_μ and $0 < \tau < 1/(1 + (2/\mu)m_\mu)$. We deduce

$$
\begin{aligned}
||l_{k+1} - l_k||_{\mathbb{C}^n} &\le ||(1-\tau)I - \tau M_\mu^{(k)}|| \, ||l_k - l_{k-1}||_{\mathbb{C}^n} \\
&\le (1-\tau)||l_k - l_{k-1}||_{\mathbb{C}^n}
\end{aligned}
\qquad (4.38)
$$

because $||A||^2 = \rho(A^*A)$. So the sequence converges. $\qquad \square$

We have a proof of the convergency of sequence $\{l_k\}_{k\in\mathbb{N}}$ if the spectral radius of the matrix $M^{(k)}{}_\mu$ is uniformly bounded with respect to k. In the next lemma, we give an estimate of the spectral radius of $M^{(k)}{}_\mu$ with the Frobenius norm.

LEMMA 4.10. *Let $\|l_0\|_{\mathbb{C}^n} \le R$; if*

$$\mu \ge 2C_{\psi^*}R\left(\log\left(\frac{R - \|V\|_{\mathbb{C}^n}}{C'_\psi}\right) + 1\right)^{-1},\tag{4.39}$$

where $C'_\psi > 0$ verifies

$$\|\psi(F)\|_{\mathbb{C}^n} \le C'_\psi \|F\|_{L^\infty(E,\mathbb{R},\sigma)},\tag{4.40}$$

then the whole sequence is bounded by R, and the spectral radius $\rho^{(k)}_\mu$ of $M^{(k)}_\mu$ satisfies

$$\rho^{(k)}_\mu \le \exp\left(-1 + \frac{2}{\mu}C_{\psi^*}R\right)\sqrt{\sum_{1\le i,j\le n}\left|\int_E q_i(\sigma)\bar{q}_j(\sigma)d\sigma\right|^2}.\tag{4.41}$$

Proof. We prove the result by induction. We assume that there exists $k \in \mathbb{N}^*$ such that for all $j \le k$, $\|l_j\|_{\mathbb{C}^n} \le R$, then

$$\|l_{k+1}\|_{\mathbb{C}^n} \le (1 - \tau)R + \tau\left(\|V\|_{\mathbb{C}^n} + C'_\psi\exp\left(-1 + \frac{2}{\mu}C_{\psi^*}R\right)\right)\tag{4.42}$$

and condition (4.39) implies

$$\|V\|_{\mathbb{C}^n} + C'_\psi\exp\left(-1 + \frac{2}{\mu}C_{\psi^*}R\right) \le R.\tag{4.43}$$

For all $k \in \mathbb{N}$, we have $\|l_k\| \le R$ and (4.41) is a direct consequence of (4.35). $\square$

PROPOSITION 4.11. *If (μ, R) satisfies condition (4.39), then γ_μ has a unique fixed point λ in the closed ball $B(0, R)$. Moreover, if $\|l_0\|_{\mathbb{C}^n} \le R$ and τ is small enough, then λ is the limit of the sequence $\{l_k\}_{k\in\mathbb{N}}$.*

We now give a more precise description of the algorithm defined by (4.36).

Algorithm 4.12.
   (1) Initialization. Given $V \in \mathbb{C}^n$, choose $l_0 \in \mathbb{C}^n$, $\mu > 0$, $\epsilon > 0$, $\tau \in]0,1]$.
   (2) Iteration k. (a) Compute

$$\gamma_\mu(l_{k-1}) = V - \int_E q(\sigma)\exp\left(-1 + \frac{2}{\mu}\psi^*[l_{k-1}](\sigma)\right)d\sigma.\tag{4.44}$$

      (b) Compute $l_k = l_{k-1} - \tau[l_{k-1} - \gamma_\mu(l_{k-1})]$.
   (3) Stopping criterion. If $|l_k - l_{k-1}| < \epsilon$, then STOP, else $k := k + 1$ and go to (2).

The algorithm converges if the regularization parameter is great enough. The main advantage of this method is that it determines a vector of $\mathbb{C}^n$ which is the fixed point of

a functional. Moreover, we have an analytic expression for this solution. The convergence of the algorithm is linear since we have shown that $\|l_{k+1} - l_k\|_{\mathbb{C}^n} \leq (1-\tau)\|l_k - l_{k-1}\|_{\mathbb{C}^n}$ in Lemma 4.9. The number τ has to be chosen as great as possible for a faster convergence.

Now we perform a sensitivity analysis of the optimal value function with respect to μ and V. Let $V \in \mathbb{C}^n$ be fixed. We suppose that for the data V, the sequence (4.31) converges for all $\mu \geq \mu_o > 0$. We define the function $\mathfrak{C}_1$ by

$$\mathfrak{C}_1 : [\mu_o, +\infty[\longrightarrow \mathbb{R}, \qquad \mu \longmapsto J_\mu(F_{\mu,V}) + \mu e^{-1}. \tag{4.45}$$

Similarly for any $\lambda > 0$ fixed, we suppose that the sequence (4.31) converges for all $V \in B(0,R)$, where $R > 0$, and we define the function $\mathfrak{C}_2$ by

$$\mathfrak{C}_2 : B(0,R) \subset \mathbb{C}^n \longrightarrow \mathbb{R}, \qquad V \longmapsto J_\mu(F_{\mu,V}) + \mu e^{-1}. \tag{4.46}$$

PROPOSITION 4.13. *Let $\mathfrak{C}_1$ and $\mathfrak{C}_2$ be the functions defined by (4.45) and (4.46).*
  (1) *$\mathfrak{C}_1$ is continuous and increasing.*
  (2) *$\mathfrak{C}_2$ is continuous and verifies*

$$\mathfrak{C}_2 \leq \|V\|_{\mathbb{C}^n}^2 + \frac{\mu}{e}. \tag{4.47}$$

Proof. Let $\mu_1 > \mu_2 \geq \mu_o$. For all $F \in K$, $J_{\mu_1}(F) > J_{\mu_2}(F)$ since $(H(F) + e^{-1}) \geq 0$, hence $\mathfrak{C}_1$ is increasing. To prove continuity, we suppose there exists a sequence $\{\mu_k\}_{k \in \mathbb{N}}$ and a $\delta > 0$ such that $\mu_k \to \mu$ and for all k, $|\mathfrak{C}_1(\mu_k) - \mathfrak{C}_1(\mu)| > \delta$. Let $\varepsilon > 0$ be small enough, so there exists $N \in \mathbb{N}$ such that $|\mu - \mu_N| < \delta(\mu - \varepsilon)/2\mathfrak{C}_1(\mu)$ and $|\mu - \mu_k| < \varepsilon$. Assume that $\mu_N < \mu$, then

$$J_\mu(F_{\mu_N,V}) - \mathfrak{C}_1(\mu_N) = (\mu - \mu_N)H(F_{\mu_N,V}) \tag{4.48}$$

so $|J_\mu(F_{\mu_N,V}) - \mathfrak{C}_1(\mu_N)| < \delta/2$ since $H(F_{\mu_N,V}) \leq \mathfrak{C}_1(\mu)/\mu_N$. Hence, $J_\mu(F_{\mu_N,V}) < \mathfrak{C}_1(\mu)$. This contradiction proves the result. The case $\mu_N \geq \mu$ can be shown similarly.

We can show the continuity of $\mathfrak{C}_2$ by the same way. The inequality (4.47) is obtained by taking $F = 0$. $\qquad\square$

We just proved the fixed point existence if μ is large enough; in fact, it is true for any $\mu > 0$. We use a scaling method to prove it.

PROPOSITION 4.14. *For every $\mu > 0$, the solution F_μ of problem $(\mathcal{P}_\mu)$ exists and verifies (4.13). Moreover, F_μ can be computed with the sequence $\{\ell'_k\}_{n \in \mathbb{N}}$ defined by*

$$\ell'_o \in \mathbb{C}^n,$$
$$\ell'_{k+1} = \left(1 - \frac{\tau_o}{\alpha}\right)\ell'_k + \tau_o \gamma_{\mu_o}(\ell'_k) \tag{4.49}$$

for some $\mu_o > \mu > 0$, $\tau_o \in\ ]0,1[$ and $\alpha = \mu_o/\mu$.

Proof. With Proposition 4.11, we know that there exist $\ell_o \in \mathbb{C}^n$, $\mu_o > 0$ and $0 < \tau_o < 1$ such that the sequence $\{\ell_k\}_{k \in \mathbb{N}}$ defined by (4.36) converges. Then the solution F_μ of problem $(\mathcal{P}_\mu)$ exists for every $\mu \geq \mu_o$.

Assume now that $\mu < \mu_o$ and set $\alpha := \mu_o/\mu > 1$. As shown in the proof of Lemma 4.9, the eigenvalues of the gradient of the function $\ell \mapsto (1 - \tau_o)\ell + \tau_o\gamma_{\mu_o}(\ell)$ are nonnegative, and the spectral radius is bounded by $1 - \tau_o$. As $-d\gamma_{\mu_o}$ is nonnegative and $\alpha > 1$, we deduce as before that the sequence $\{\ell'_k\}$ defined by

$$\ell'_o \in \mathbb{C}^n,$$
$$\ell'_{k+1} = \left(1 - \frac{\tau_o}{\alpha}\right)\ell'_k + \tau_o\gamma_{\mu_o}(\ell'_k) \tag{4.50}$$

is converging to some ℓ' that verifies

$$\ell' = \left(1 - \frac{\tau_o}{\alpha}\right)\ell' + \tau_o\gamma_{\mu_o}(\ell'),$$
$$\frac{\tau_o}{\alpha}\ell' = \tau_o\gamma_{\mu_o}(\ell'). \tag{4.51}$$

Setting $\lambda = \ell'/\alpha$ yields

$$\lambda = \gamma_{\mu_o}(\alpha\lambda) = V - \psi\left[\exp\left(-1 + \frac{2}{\mu_o}\psi^*(\alpha\lambda)\right)\right],$$
$$\lambda = V - \psi\left[\exp\left(-1 + \frac{2\alpha}{\mu_o}\psi^*(\lambda)\right)\right] = V - \psi[\exp(-1 + 2\mu\psi^*(\lambda))] = \gamma_\mu(\lambda). \tag{4.52}$$

We conclude with Lemma 4.8 that $(\mathscr{P}_\mu)$ has a (unique) solution. $\qquad\square$

In summary, we have a finite-dimensional algorithm that converges to the solution of maximum entropy regularization for all $\mu > 0$. However, Proposition 4.14 shows that the number of iteration increases when μ decreases since the speed of convergency is $(1 - \tau_o\mu/\mu_o)$.

We know the analytic expression of the solution and the problem is now a finite-dimensional one. However, the entropy involved in the functional we have minimized is not the entropy in the sense of Definition 3.1 because the solution is not a probability density. The algorithm we present in the next section allows to find a solution which maximize the true entropy. It is a variant of the previous one.

5. Computing the probability density

We look for a solution that maximizes entropy in the sense of Definition 3.1. This leads to the problem

$$(\widetilde{\mathscr{P}}_\mu)$$

$$\min \|V - \alpha\psi[F]\|^2_{\mathbb{C}^n} + \mu H(F),$$
$$(\alpha, F) \in \mathbb{R}^+ \times \{f \in K \mid \|f\|_{L^1(E,\mathbb{R},\sigma)} = 1\}. \tag{5.1}$$

Here, $\alpha \in \mathbb{R}^+$ is the total power and F is a probability density because of the constraint $F \in \{f \in K \mid \|f\|_{L^1(E,\mathbb{R},\sigma)} = 1\}$. So, if the solution exists, it minimizes negentropy as defined in Definition 3.1.

We cannot solve (5.1) in the same way as problem $(\mathcal{P}_\mu)$. Indeed, the cost functional is no more convex because of the bilinear term $\alpha\psi[F]$. On the other hand, the feasible domain is convex since for all $f,g \in \{f \in K \mid \|f\|_{L^1(E,\mathbb{R},\sigma)} = 1\}$, for all $\lambda \in [0,1]$, we have $\lambda f + (1-\lambda)g \in K$, and

$$\|\lambda f + (1-\lambda)g\|_{L^1(E,\mathbb{R},\sigma)} = \lambda\|f\|_{L^1(E,\mathbb{R},\sigma)} + (1-\lambda)\|g\|_{L^1(E,\mathbb{R},\sigma)} = 1 \tag{5.2}$$

by positivity of f, g. Anyway, if F is solution of $(\mathcal{P}_\mu)$, we can view $\|F\|_{L^1(E,\mathbb{R},\sigma)}$ as an approximation of the power. So we use the previous section results on $(\mathcal{P}_\mu)$. We define the sequence $\{l_k\}_{k\in\mathbb{N}}$ by

$$\begin{aligned} l_o &\in \mathbb{C}^n, \\ \delta_{k,\mu} &= \frac{\mathrm{Re}\left(\langle V, g_\mu(l_k)\rangle_{\mathbb{C}^n}\right)}{\langle g_\mu(l_k), g_\mu(l_k)\rangle_{\mathbb{C}^n}}, \\ l_{k+1} &= V - \delta_{k,\mu} g_\mu(l_k), \end{aligned} \tag{5.3}$$

where $g_\mu : \mathbb{C}^n \to \mathbb{C}^n$ is defined by

$$g_\mu(l) = V - \gamma_\mu(l) = \int_E q(\sigma)\exp\left(-1 + \frac{2}{\mu}\psi^*[l](\sigma)\right)d\sigma. \tag{5.4}$$

We assume that

$$\forall l \in \mathbb{C}^n, \quad \mathrm{Re}\left(\langle V, g_\mu(l)\rangle_{\mathbb{C}^n}\right) \geq 0; \tag{5.5}$$

since $\delta_{k,\mu}$ is an estimation of the power, it has to be positive. If $\delta_{k,\mu} \to \delta^*$ and $l_k \to l^*$, then

$$l^* = V - \delta^* g_\mu(l^*). \tag{5.6}$$

Last equation is quite similar to the necessary and sufficient condition of optimality (4.9) of problem $(\mathcal{P}_\mu)$.

PROPOSITION 5.1. *Let $l^* = V - \delta^* g_\mu(l^*)$ with $\delta^* > 0$; define*

$$F_{l^*} = \exp\left(-1 + \frac{2}{\mu}\psi^*[l^*]\right), \qquad f_{l^*} = \frac{F_{l^*}}{\|F_{l^*}\|_{L^1(E,\mathbb{R},\sigma)}}. \tag{5.7}$$

Then f_{l^*} is a probability density and it is the unique solution of problem

$$\min\left\|\frac{V}{\delta^*\|F_{l^*}\|_{L^1(E,\mathbb{R},\sigma)}} - \psi(F)\right\|^2_{\mathbb{C}^n} + \frac{\mu}{\delta^*\|F_{l^*}\|_{L^1(E,\mathbb{R},\sigma)}}\int_E \phi \circ F(\sigma)d\sigma, \tag{5.8}$$
$$F \in \{f \in \mathbb{H} \mid f \geq 0 \ \sigma\text{-a.e.}, \ \|f\|_{L^1(E,\mathbb{R},\sigma)} = 1\}.$$

Moreover, f_{l^*} verifies

$$\forall \alpha \in \mathbb{R}, \quad \|V - \delta^*\|F_{l^*}\|_{L^1(E,\mathbb{R},\sigma)}\psi[f_{l^*}]\|^2_{\mathbb{C}^n} \leq \|V - \alpha\psi[f_{l^*}]\|^2_{\mathbb{C}^n}. \tag{5.9}$$

Proof. We have

$$l^* = V - \delta^* g_\mu(l^*) \implies -2\psi^*[V - \delta^* g_\mu(l^*) - l^*] = 0$$

$$\implies -2\psi^*[V - \delta^* g_\mu(l^*)] + \mu\left(\frac{2}{\mu}\psi^*[l^*]\right) = 0 \qquad (5.10)$$

$$\implies -2\psi^*[V - \delta^*\psi[F_{l*}]] + \mu(1 + \ln F_{l*}) = 0.$$

Dividing the last equation by δ^* gives a sufficient optimality condition for problem (5.8). So F_{l*} is the unique solution to this problem. As $\{f \in \mathbb{H} \mid f \geq 0 \ \sigma\text{-a.e.}, \ \|f\|_{L^1} = 1\} \subset K$, we get the conclusion.

To show (5.9), we suppose that $\psi[f_{l*}] \neq 0$ (otherwise the result is obvious). We define

$$\theta : \mathbb{R} \longrightarrow \mathbb{R}, \qquad \alpha \longmapsto \|V - \alpha\|F_{l*}\|_{L^1(E,\mathbb{R},\sigma)}\psi[f_{l*}]\|_{\mathbb{C}^n}^2. \qquad (5.11)$$

It is a continuous and strictly convex function. The unique minimizer of θ (denoted α^*) verifies

$$\theta'(\alpha) = -2\|F_{l*}\|_{L^1(E,\mathbb{R},\sigma)} \mathrm{Re}\left(\langle V - \alpha^*\|F_{l*}\|_{L^1(E,\mathbb{R},\sigma)}\psi[f_{l*}], \psi[f_{l*}]\rangle_{\mathbb{C}^n}\right) = 0, \qquad (5.12)$$

that is,

$$\alpha^* = \frac{\mathrm{Re}\left(\langle V, \psi[f_{l*}]\rangle_{\mathbb{C}^n}\right)}{\|F_{l*}\|_{L^1(E,\mathbb{R},\sigma)}\|\psi[f_{l*}]\|_{\mathbb{C}^n}^2} = \delta^*. \qquad (5.13)$$

□

The above proposition shows that if the sequence given by (5.3) converges, then its limit provides a power density $(\delta^*, f_{l*} d\sigma)$, which minimizes the error with the data V. Moreover, f_{l*} has a minimal negentropy in the probabilistic sense. Now, we show that the obtained solution is the unique solution of problem $(\widetilde{\mathcal{P}}_{\delta^*\|F_{l*}\|_{\mu L^1(E,\mathbb{R},\sigma)}})$. We call

$$\mathfrak{J}_\mu(\alpha, F) = \|V - \alpha\psi[F]\|_{\mathbb{C}^n}^2 + \mu H(F), \qquad (5.14)$$

we have $\mathfrak{J}_\mu(1, F) = J_\mu(F)$.

The previous result proves that if the sequence defined by (5.3) converges, then its limit is a solution of problem $(\widetilde{\mathcal{P}}_{\delta^*\|F_{l*}\|_{\mu L^1(E,\mathbb{R},\sigma)}})$. We have a result of convergence.

PROPOSITION 5.2. *Assume that there exists $c > 0$ such that*

$$\forall \mu \geq 0, \ \forall l \in \mathbb{C}^n, \quad \|g_\mu(l)\|_{\mathbb{C}^n} \geq c, \qquad (5.15)$$

then there exists a μ_o such that the sequence defined by (5.3) converges for every $\mu \geq \mu_o$ large enough.

Proof. We have to compute the derivative of the function defining the sequence (5.3):

$$\varphi_\mu : l \longmapsto V - \frac{\mathrm{Re}\left(\langle V, g_\mu(l)\rangle_{\mathbb{C}^n}\right)}{\langle g_\mu(l), g_\mu(l)\rangle_{\mathbb{C}^n}} g_\mu(l). \qquad (5.16)$$

We get

$$
\varphi_\mu'(l) = - \frac{\mathrm{Re}\left(\langle V, g_\mu(l)\rangle_{\mathbb{C}^n}\right)}{\|g_\mu(l)\|_{\mathbb{C}^n}^2} \nabla g_\mu(l)
$$
$$
- g_\mu(l)\left[\frac{\mathrm{Re}\left(\langle V, \nabla g_\mu(l)\rangle_{\mathbb{C}^n}\right)}{\|g_\mu(l)\|_{\mathbb{C}^n}^2} - 2\frac{\mathrm{Re}\left(\langle V, \nabla g_\mu(l)\rangle_{\mathbb{C}^n}\right)}{\|g_\mu(l)\|_{\mathbb{C}^n}^2}\frac{\mathrm{Re}\left(\langle g_\mu(l), \nabla g_\mu(l)\rangle_{\mathbb{C}^n}\right)}{\|g_\mu(l)\|_{\mathbb{C}^n}^2}\right].
$$

$$(5.17)$$

A short computation gives

$$
\sup_{l\in\mathbb{C}^n}\|\varphi_\mu'(l)\| \le 2\|V\|\kappa_\mu(1 + 2\kappa_\mu), \tag{5.18}
$$

where

$$
\kappa_\mu \overset{\text{def}}{=} \sup_{l\in\mathbb{C}^n}\frac{\|\nabla g_\mu(l)\|}{\|g_\mu(l)\|}. \tag{5.19}
$$

Thanks to assumption (5.15) and (4.33), there exists a constant $c' > 0$ such that $\kappa_\mu < c'/\mu$. It is clear that there exists $\mu_o > 0$ such that φ_μ is contractive for every $\mu \ge \mu_o$. □

Remark 5.3. Assumption (5.15) seems to be surprising, but in fact it is primordial. This assumption can easily be derived from the physical model: $g_\mu(l)$ is the spectral matrix corresponding to the WDF $\exp(-1 + (2/\mu)\psi^*[l]) > 0$. Since the WDF is positive, it means that the electromagnetic wave has a nonzero energy, and thus the spectral matrix cannot be zero. We can see (5.15) from the model noting that the spectral matrix trace is the constant function 1 on the sphere.

Remark 5.4. We have not prove the convergency for all $\mu > 0$. Anyway, in practice, $V \ll \psi$. So we have to perform a (physical) normalization process; this means that we use the operator $\varepsilon\psi$ instead of ψ, where $\varepsilon > 0$ is a small number (usually $\varepsilon \simeq \|V\|/\|\psi\|$). Hence, we have to solve the problem $(\widetilde{\mathscr{P}}_{\mu/\varepsilon^2})$ (according to the proof of Proposition 4.14), and the new regularization parameter $\tilde{\mu} = \mu/\varepsilon^2$ is large.

We now can write the algorithm more precisely and we introduce a relaxation parameter τ useful for numerical computations.

Algorithm 5.5.
  (1) Initialization. Given $V \in \mathbb{C}^n$, choose $\ell_o \in \mathbb{C}^n$, $\mu > 0$, $\tau \in]0, 1[$, $\epsilon > 0$.
  (2) Iteration k. (a) Compute

$$
g_\mu(\ell_{k-1}) = \int_E q(\sigma)\exp\left(-1 + \frac{2}{\mu}\psi^*[\ell_{k-1}](\sigma)\right)d\sigma. \tag{5.20}
$$

  (b) Compute

$$
\delta_{k-1,\mu} = \frac{\mathrm{Re}\left(\langle V, g_\mu(\ell_{k-1})\rangle_{\mathbb{C}^n}\right)}{\langle g_\mu(\ell_{k-1}), g_\mu(\ell_{k-1})\rangle_{\mathbb{C}^n}}. \tag{5.21}
$$

  (c) Compute $\ell_k = (1 - \tau)l_{k-1} + \tau[V - \delta_{k-1,\mu}g_\mu(\ell_{k-1})]$.

(3) Stopping criterion. If $|\ell_k - \ell_{k-1}| < \epsilon$, then STOP, else $k := k + 1$ and go to (2).

Remark 5.6. In this section, we have supposed that $R(\psi^*) \subset L^\infty(E, \mathbb{R}, \sigma)$ to establish the algorithms. If this condition is not verified, we can slightly modify the operator ψ to overcome this problem. More precisely, if $\psi : F \mapsto \int_E q(\sigma) F(\sigma) d\sigma$ with $q \notin L^\infty(E, \mathbb{C}^n, \sigma)$, we can find a $q^\varepsilon \in L^\infty(E, \mathbb{C}^n, \sigma)$ such that $\|q - q^\varepsilon\|_{L^2(E, \mathbb{C}^n, \sigma)} \leq \varepsilon$. The function q^ε allows to define a new operator $\psi_\varepsilon : F \mapsto \int_E q^\varepsilon(\sigma) f(\sigma) d\sigma$, which verifies $R(\psi_\varepsilon) \subset L^\infty(E, \mathbb{R}, \sigma)$ and $\|\psi_\varepsilon[F] - \psi[F]\|_{\mathbb{C}^n} \leq \varepsilon \|F\|_{\mathbb{H}}$. This new operator can be used for the computation instead of ψ, and it verifies the desired condition.

6. Numerical tests

In this section, we perform some numerical tests for the resolution of (2.3) with the two algorithms previously described. For these tests, we consider the propagation of an electromagnetic wave in vacuum. We give below the corresponding expression of operator ψ. We denote by $\epsilon_r = \|V - \psi[F_{\mu,V}]\|_{\mathbb{C}^n}^2$ the error and α is the computed approximation of total power P.

For the numerical computation of the solution, we need to find $\tau > 0$ small enough for the algorithms to converge. The number τ must be great enough for a fast computation. To determine this parameter, we have used the following rules:

    (i) let $\tau_o > 0$, $k = 0$, and $\ell_o = 0$;
    (ii) iteration $k : \tau_k$ and ℓ_k are known; compute ℓ_{k+1};
    (iii) if $\|\ell_{k+1} - \ell_k\|_{\mathbb{C}^n} \geq \|\ell_k - \ell_{k-1}\|_{\mathbb{C}^n}$, then $\tau_{k+1} = \tau_k r$, where $0 < r < 1$.

In the computation, we used $r = 0.5$. We can also utilize a linear search strategy for the best value of τ. We have chosen to stop the algorithms when $\|l_{k+1} - l_k\|_{\mathbb{C}^n} < \epsilon$, where $\epsilon > 0$ is chosen small enough (here $\epsilon = 10^{-12}$).

All numerical tests were performed on a Power Mac $2 \times$ G4 with the scientific software Scilab 2.6. For these examples, the average time for one iteration was 28.5 milliseconds. This is corresponding to the average number of 35 iterations per second.

6.1. Wave distribution function in vacuum.
In the case of an electromagnetic wave propagating in vacuum, we have the relation

$$V = \int_0^\pi \int_0^{2\pi} q(\theta, \phi) F(\theta, \phi) \sin\theta \, d\phi \, d\theta, \tag{6.1}$$

where V is the data vector, F is the WDF of the electromagnetic wave, and q is the integrating kernel of vacuum. Writing the components of the spectral matrix as a vector provides the data V. We see with (6.1) that we integrate over the unit sphere: $\theta \in [0, \pi]$ denotes the polar angle and $\phi \in [0, 2\pi]$ the azimuthal angle. This is quite clear because WDF is the directional distribution of the power of electromagnetic wave. The expression of q is analytically known [7].

According to (6.1), we put $E = [0, \pi] \times [0, 2\pi]$, $d\sigma = (\sin\theta/2\pi) d\theta d\phi$, and we define the operator ψ by

$$\psi : \mathbb{H} \longrightarrow \mathbb{C}^n, \qquad F \longmapsto \int_0^\pi \int_0^{2\pi} q(\theta, \phi) F(\theta, \phi) \sin\theta \, d\phi \, d\theta. \tag{6.2}$$
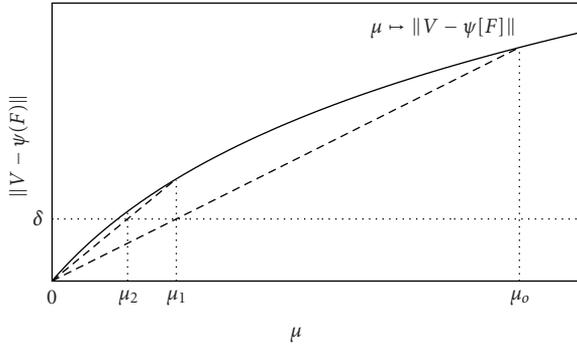
Figure 6.1. Iterative search of the regularization parameter μ that fits the Morozov discrepancy principle.

This operator is continuous on $\mathbb{H}$ and the adjoint operator verifies $R(\psi^*) \subset L^\infty(E, \mathbb{C}^n, \sigma)$ [7]. From the expression of ψ, we see that we have to compute a double integral. For this computation, we have used the Gauss-Legendre quadrature method with 20 points on the interval $[0, \pi]$ and 40 points on $[0, 2\pi]$. In the following tests, this integration method seems to be accurate enough because the results do not change dramatically if the number of Gauss points is larger.

Numerical validation of algorithms was done with simulated data. In the first tests, V was obtained for a WDF that was a sum of Dirac measures. The data vector of a Dirac in (θ_d, ϕ_d) of power P is simulated by taking $V = P/\sin(\theta_d)q(\theta_d, \phi_d)$. The case of a Dirac is important because it physically corresponds to an electromagnetic plane wave. Second tests are performed with data corresponding to an a.e. continuous WDF. The simulation has been made by computing the integral (6.1). For some of these examples, a noisy data V^δ was introduced with $\|V - V^\delta\|_{\mathbb{C}^n} \leq \delta$, where $\delta > 0$ is the noise level.

The choice of the regularization parameter is a difficult task in the case of noisy data. It must be chosen small enough to make a small error, and large enough to guarantee stability. A commonly used rule is the discrepancy principle of Morozov [4]. We choose μ such that $\|V - \psi[F_\mu]\|_{\mathbb{C}_n} = \delta$, that is, we do not make a smaller error than the noise level. In the numerical tests, since we used simulated data, we decided to choose μ such that $\delta \leq \|V - \psi[F_\mu]\|_{\mathbb{C}_n} \leq 2\delta$. This choice is made *a posteriori* using the fact that the function $\mu \mapsto \|V - \psi[F_{\mu,V}]\|_{\mathbb{C}^n}$ is quite linear when μ is small (see Figure 6.1). Note that the noise level δ is known from the experience.

6.2. Test of Algorithm 4.12.
We first used Algorithm 4.12 for data corresponding to a Dirac measure or a sum of Dirac measures. We present the results in the case of one-single Dirac and the sum of three Dirac measures to illustrate the behaviour of the algorithm that is able to detect more than one direction.

6.2.1. Case of one Dirac measure. We have made computations with $\theta_d = 1.5$, $\phi_d = 2$ and for $P_i = i/2$ with $i = 1, \ldots, 50$. For each value of P_i, we have computed the relative power

error given by

$$\epsilon_\alpha^i = \frac{|\alpha_i - P_i|}{P_i}, \tag{6.3}$$

where α_i is the computed approximation of total power P_i. For this example, we set $\mu = 1$. The function $\epsilon_\alpha(P)$ is plotted in Figure 6.1. We see that the error increases when P decreases; it is less than 10% for a power $P \geq 4$. We have also computed the mean direction $(\bar{\theta}, \bar{\phi})$ given by

$$\begin{aligned}
\bar{\theta} &= \frac{1}{\|F\|_{L^1(E, \mathbb{R}, \sigma)}} \int_E \theta F(\theta, \phi) d\sigma(\theta, \phi), \\
\bar{\phi} &= \frac{1}{\|F\|_{L^1(E, \mathbb{R}, \sigma)}} \int_E \phi F(\theta, \phi) d\sigma(\theta, \phi)
\end{aligned} \tag{6.4}$$

of a Dirac with $\phi_d = 2$, $P = 8$ and for θ_d varying from 0 to π. From these values, we can compute the angles errors ϵ_θ and ϵ_ϕ by

$$\epsilon_\theta = \frac{|\theta_d - \bar{\theta}|}{\pi}, \qquad \epsilon_\phi = \frac{|\phi_d - \bar{\phi}|}{2\pi}. \tag{6.5}$$

We can see how the average direction changes with θ_d. We see that the average direction is close to the true one when it is far enough from the poles. The "large" error on ϕ_d near the poles can be explained by the "bad" representation of the unit sphere in (6.4). Similarly, we obtain a large error if ϕ_d is chosen near 0 or 2π. Another value for ϕ_d far enough from the edge of E gives a similar result as that plotted in Figure 6.2. We conclude that the results obtained in this case are satisfactory. The method is able to detect the direction and the power of a Dirac with small errors if the direction is not too close to the edge of E and if the power is great enough.

6.2.2. Case of the sum of three Dirac measures.   We have built an example for a sum of three Dirac measures which support the vertices of an equilateral triangle on E. For this example, we have added a noise to the simulated data V. It verifies $\|V - V^\delta\|_{\mathbb{C}^n} < 1$ (precisely $= 0.764$, this correspond to a relative noise level of 3.1%). The regularization parameter μ was chosen according to the noise level; we have taken $\mu = 0.5$ because the corresponding solution has a small error and $1 = \delta < \|V - \psi[F]\|_{\mathbb{C}^n}$, more precisely, the relative error is 5.5% (see Table 6.1 for the results). We plot the contour of the solution on Figure 6.3. We observe three peaks centered on the directions of the three Dirac measures. The solution is in this way satisfactory.

   We have performed a sensitivity analysis with respect to μ. Indeed, in Section 3, we gave an interpretation of entropy as the "distance" to the noninformative probability density g. To see the numerical effects of entropy, we computed the solution for the same data with different values for μ. Theoretically, a great value of μ leads to a solution with a great entropy, that is a solution not too "far" from g. In Figure 6.3, we see the result obtain for $\mu = 50$ and the different results are reported in Table 6.1. We see that the error ϵ_r quickly increases with μ while H is decreasing. The number of iteration is dramatically increasing when μ becomes small, so that we are not able to compute the μ-solution for all μ with
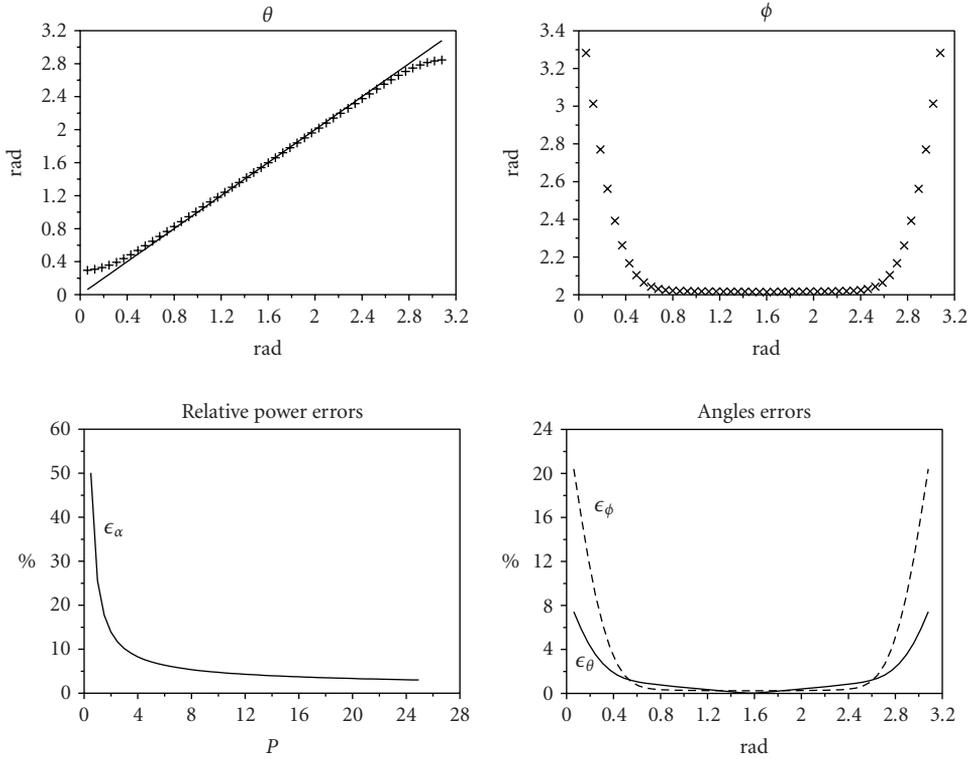
Figure 6.2. Variation of the average direction $\bar{\theta}$ (plotted with +, the true value is the solid line), $\bar{\phi}$ (plotted with ×), and errors ϵ_θ, ϵ_ϕ for a Dirac with $P = 8$, $\phi_d = 2$, and θ_d varying from 0 to π. The plot named "relative power error" show the variation of ϵ_α for a Dirac with $\theta_d = 1.5$, $\phi_d = 2$ and P varying from 0.5 to 25.

Table 6.1. Results obtained for a simulated noisy data V of a sum of three Dirac measures. We have made four computations for different values of the regularization parameter μ. The Min and Max are, respectively, the minimum and the maximum of the solution on E.

μ	α	ϵ_α	ϵ_r	H	J_μ	Min	Max	#It.	τ
0.5	9.891	2.07%	1.842	-1.452	5.741	$2.200E-6$	6.947	8377	$2.5E-3$
1	9.868	1.83%	3.159	-1.660	9.360	$1.521E-4$	4.597	3317	$6.25E-2$
5	9.814	1.27%	10.27	-1.990	24.691	$2.678E-2$	3.136	465	0.05
50	8.008	17.35%	117.633	-2.338	17.862	0.195	1.725	7	1

this algorithm. We observe that we need more and more iteration when $\mu \to 0$ to converge. The solution obtained for $\mu = 50$ is very flat and has a very large error $\epsilon_r = 117.6$. We remark that the optimal value function is not increasing but there is no contradiction with Proposition 4.13 since we have to add μ/e to obtain an increasing function.
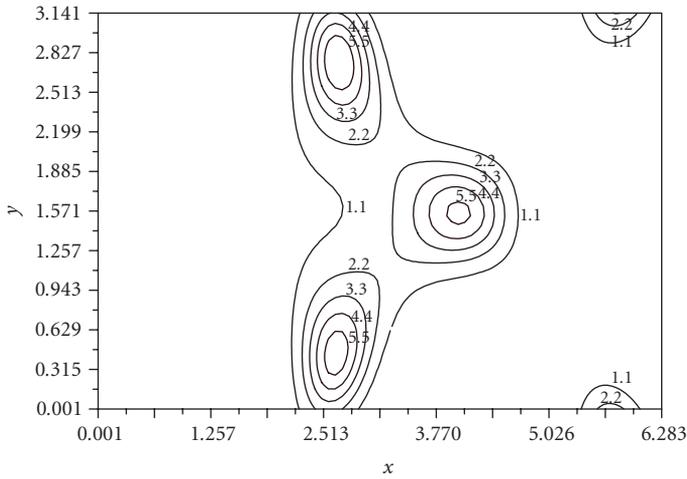
(a)
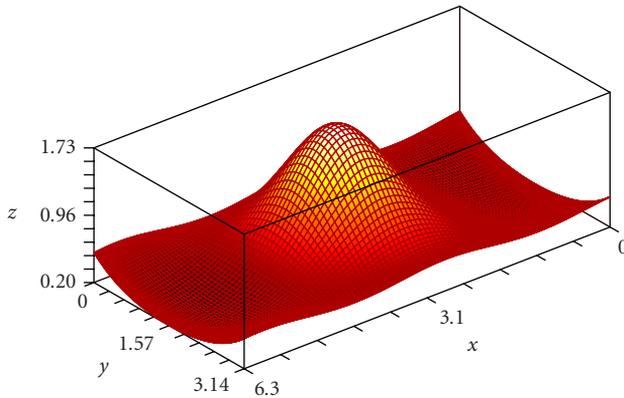
(b)

Figure 6.3. (a) Contour plot of the solution obtained by algorithm (4.36) for three Dirac measures with $\mu = 0.5$. (b) Solution obtained for three Dirac measures with $\mu = 50$.

6.3. Test of Algorithm 5.5 on a continuous density. In the previous section, we presented tests of the first algorithm to identify a three-Dirac distribution. Now we give an example of reconstruction of an a.e. continuous distribution with Algorithm 5.5. Let $F_t(\theta, \phi) = 4(\cos^2 \theta + \sin^2 \phi)$, the function used to simulate the data vector V. In this example, we use a noisy data V^δ with $\| V - V^\delta \| = 0.670$, which is about 5.1% compared with $\| V \|_{\mathbb{C}^n}$. To avoid so-called inverse crime, we have used a finer quadrature rule to compute

the data V: more precisely, we used 64 points on the interval $[0, \pi]$ and 128 points on $[0, 2\pi]$.

The function F_t is represented in Figure 6.4. We use a spherical plot: the value of the function is described by a gray-level code on each hemisphere. We see that F_t is not continuous at the two poles of the sphere. This discontinuity is explained by the "bad" representation of the unit sphere: $E = [0, \pi] \times [0, 2\pi]$. The solution obtained by Algorithm 5.5 is also plotted in Figure 6.4.

We see that the computed solution looks like F_t. We used the following parameters: $\epsilon = 10^{-12}$, $\mu = 3$, $\tau = 0.008$. The algorithm stopped after 1818 iterations, the error with respect to the data ϵ_r and the negentropy H are, respectively, 1.127 and -2.460. (The error is about 8.0% compared to $\|V\|_{\mathbb{C}^n}$.) We see that this solution is satisfying because ϵ_r is greater than the noise level but it is small enough anyway.

This example allows to compare problems $(\mathcal{P}_\mu)$ and (5.1). Let μ_e be defined by $\mu_e = \mu\delta^* \|F_{l*}\|_{L^1(E, \mathbb{R}, \sigma)}$ with the notations of Proposition 5.1. With this proposition, we know that f_{l*} achieves the minimum of functional

$$F \longmapsto \left\|V - \frac{\mu_e}{\mu}\psi[F]\right\|_{\mathbb{C}^n}^2 + \mu_e H(F) \tag{6.6}$$

over K. So f_{l*} is the solution of $(\widetilde{\mathcal{P}}_{\mu_e})$. We can now compare the two problems by computing the solution of $(\mathcal{P}_{\mu_e})$.

We found $\mu_e = 176.35$ with the computation of the solution of $(\widetilde{\mathcal{P}}_{\mu_e})$. The computed solution of problem $(\mathcal{P}_{\mu_e})$ is very far from the data ($\epsilon_r = 5133.6$). So, for this example, the solution of $(\widetilde{\mathcal{P}}_{\mu_e})$ is much better than the solution of $(\mathcal{P}_{\mu_e})$. For (5.1), we get a much smaller error for a much larger μ.

7. Conclusion

In plasma physics, the determination of the directional power density distribution of an electromagnetic wave from the measurement of the field components is an inverse ill-posed problem. This problem can be written as

$$\psi[F] = V, \tag{7.1}$$

where ψ is a linear bounded operator from $\mathbb{H} = L^2(E, \mathbb{R}, \sigma)$ to $\mathbb{C}^n$, V the spectral matrix, and F the wave distribution function (WDF). Lefeuvre and Delannoy [6] have proposed to solve this problem by maximizing an entropic term $-H(F) = -\int_E F\ln F \, d\sigma$ under the constraint $\psi[F] = V$. However, this constraint is too "restrictive," it indeed limits the feasible domain to a linear subspace of $\mathbb{H}$. That is why we have studied the relaxed problem $(\mathcal{P}_\mu)$:

$$\begin{aligned} \min J_\mu(F) &= \|V - \psi[F]\|_{\mathbb{C}^n}^2 + \mu H(F), \\ F &\in K = \{f \in \mathbb{H} \mid f \geq 0 \; \sigma\text{-a.e.}\}, \end{aligned} \tag{7.2}$$

where μ is a regularization parameter. The latter parameter has to be chosen small enough to allow a solution with a small error thanks to the data V, and large enough for stability.
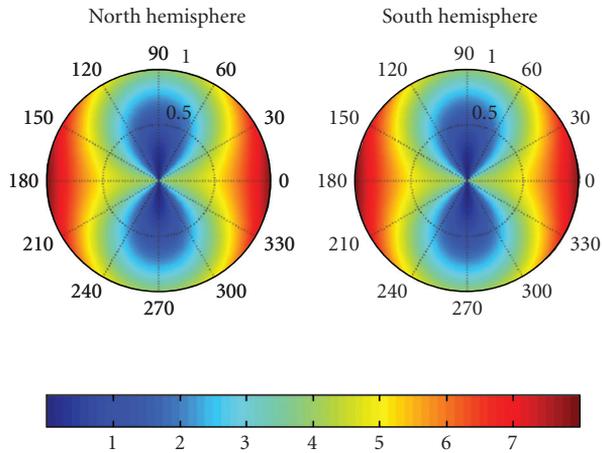
(a)
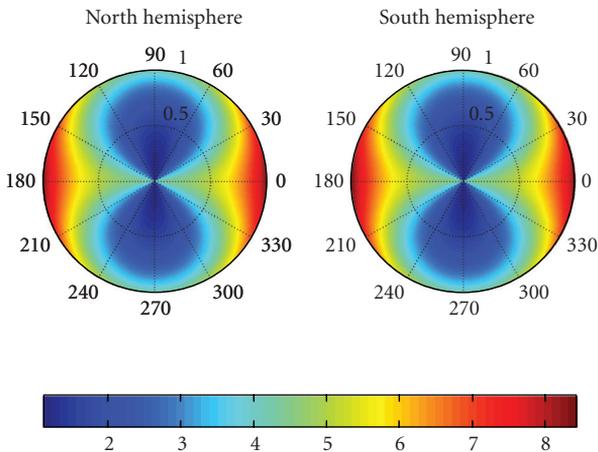
(b)

Figure 6.4. (a) Spherical plot of the initial function F_t which provides the data vector V. (b) Solution obtained by Algorithm 5.5 corresponding to the data V of the function F_t.

Solving $(\mathcal{P}_\mu)$ permits us to search the solution in a much larger domain. More precisely, the obtained solution verifies $\|\psi[F] - V\|_{\mathbb{C}^n} \le \varepsilon$ with $\varepsilon > 0$, this inequality is clearly more realistic from the numerical point of view. The regularization by the negentropy functional is also important from the physical point of view. Instead of using standard method to solve the problem $(\mathcal{P}_\mu)$, we have first built a fixed point algorithm in $\mathbb{C}^n$ (Algorithm 4.12) thanks to a sufficient condition of optimality. Moreover, the uniqueness

of the obtained solution has been proved. We have shown the existence of the solution for every value of the regularization parameter. Nevertheless, in the numerical computation, we have seen that it is not possible to compute the solution for too small value since the computational time goes to infinity.

As both the above method and the Lefeuvre's one give solutions which do not maximize entropy in the probabilistic sense, since they are not probability densities, we have built a second algorithm (Algorithm 5.5) derived from the first one which allows the "true" entropy to be maximized.

We have performed numerical tests on experimental data from a satellite that is devoted to the study of magnetosphere. We have compared the solution given by the algorithms described in the present paper to the ones obtained with two different methods that are commonly used by physicists. Results show that the method we described is more accurate and stable. They are reported in [9].

Acknowledgment

References

[1]   U. Amato and W. Hughes, *Maximum entropy regularization of Fredholm integral equations of the first kind*, Inverse Problems **7** (1991), no. 6, 793–808.

[2]   H. Brézis, *Analyse Fonctionnelle. Théorie et Application*, Dunod, Paris, 1999.

[3]   H. W. Engl and G. Landl, *Convergence rates for maximum entropy regularization*, SIAM J. Numer. Anal. **30** (1993), no. 5, 1509–1536.

[4]   A. Kirsch, *An Introduction to the Mathematical Theory of Inverse Problems*, Applied Mathematical Sciences, vol. 120, Springer-Verlag, New York, 1996.

[5]   F. Lefeuvre, *Analyse de champs d'ondes électromagnétiques aléatoires observées dans la magnétosphère à partir de la mesure simultanée de leurs six composantes*, Ph.D. thesis, University of Orléans, Orléans, 1977.

[6]   F. Lefeuvre and C. Delannoy, *Analysis of random electromagnetic wave field by a maximum entropy method*, Ann. Telecommun. **34** (1979), 204–213.

[7]   F. Lefeuvre, D. Lagoutte, and M. Menvielle, *On use of the wave distribution concept to determine the directions of arrivals of radar echoes*, Planetary and Space Science **48** (2000), no. 1–14, 1321–1328.

[8]   A. S. Leonov, *Generalization of the maximal entropy method for solving ill-posed problems*, Siberian. Math. J. **41** (2000), no. 4, 716–721.

[9]   O. Prot, O. Santolik, and J. G. Trotignon, *An entropy regularization method applied to an ELF Hiss event wave distribution function identification*, submitted to Journal of Geophysical Research.

[10] J. E. Shore and R. W. Johnson, *Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy*, IEEE Trans. Inform. Theory **26** (1980), no. 1, 26–37.

[11] A. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill-Posed Problems*, John Wiley & Sons, New York, 1977.

[12] A. N. Tikhonov, A. V. Goncharsky, V. V. Stepanov, and A. G. Yagola, *Numerical Methods for the Solution of Ill-Posed Problems*, Mathematics and Its Applications, vol. 328, Kluwer Academic Publishers, Dordrecht, 1995.

Olivier Prot: UMR 6115, LPCE/CNRS, 3A Avenue de la Recherche Scientifique, 45071 Orléans Cedex 2, France

Current address: UMR 6628, MAPMO/CNRS, Université d'Orléans, BP 6759, 45067 Orléans Cedex 2, France

E-mail address: prot@cnrs-orleans.fr

Maïtine Bergounioux: UMR 6628, MAPMO/CNRS, Université d'Orléans, BP 6759, 45067 Orléans Cedex 2, France

E-mail address: maitine.bergounioux@univ-orleans.fr

Jean Gabriel Trotignon: UMR 6115, LPCE/CNRS, 3A Avenue de la Recherche Scientifique, 45071 Orléans Cedex 2, France