

## ON THE CURVED EXPONENTIAL FAMILY IN THE STOCHASTIC APPROXIMATION EXPECTATION MAXIMIZATION ALGORITHM

VIANNEY DEBAVELAERE<sup>1,\*</sup> AND STÉPHANIE ALLASSONNIÈRE<sup>2</sup>

**Abstract.** The Expectation-Maximization Algorithm (EM) is a widely used method allowing to estimate the maximum likelihood of models involving latent variables. When the Expectation step cannot be computed easily, one can use stochastic versions of the EM such as the Stochastic Approximation EM. This algorithm, however, has the drawback to require the joint likelihood to belong to the curved exponential family. To overcome this problem, [16] introduced a rewriting of the model which “exponentializes” it by considering the parameter as an additional latent variable following a Normal distribution centered on the newly defined parameters and with fixed variance. The likelihood of this new exponentialized model now belongs to the curved exponential family. Although often used, there is no guarantee that the estimated mean is close to the maximum likelihood estimate of the initial model. In this paper, we quantify the error done in this estimation while considering the exponentialized model instead of the initial one. By verifying those results on an example, we see that a trade-off must be made between the speed of convergence and the tolerated error. Finally, we propose a new algorithm allowing a better estimation of the parameter in a reasonable computation time to reduce the bias.

**Mathematics Subject Classification.** 62F12, 62L20.

Received February 9, 2021. Accepted September 14, 2021.

### 1. INTRODUCTION

With the increase of data, parametric statistical models have become a crucial tool for data analysis and understanding. To be able to describe complex natural phenomena (epidemiology, ecology, finance, disease evolution, etc.), the models have an increasing complexity. Some of them are based on observed features or data which are assumed to be generated from a latent random effect. A usual example is the family of mixed effects models which have been used in pharmacokinetic, pharmacodynamic, shape analysis, etc. In such a context, one aims at optimizing the model parameter to maximize the likelihood of the observed dataset. This likelihood is also called the incomplete one as the latent variables are unknown.

Formally, this writes as follow: let  $y \in \mathbb{R}^n$  be the observation and  $\theta \in \Theta$  the model parameter. We call  $g$  the incomplete likelihood:

$$g(y, \theta) = \int_Z f(y, z, \theta) dz.$$

---

*Keywords and phrases:* Expectation maximization, curved exponential, mixed effect models, convergence analysis.

<sup>1</sup> Centre de Mathématiques Appliquées, École Polytechnique, Palaiseau, France.

<sup>2</sup> Centre de Recherche des Cordeliers, Université Paris Descartes, Paris, France.

\* Corresponding author: [vianney.debavelaere@polytechnique.edu](mailto:vianney.debavelaere@polytechnique.edu)

In that case,  $z$  is the latent or missing variable and  $f$  is the joint likelihood of the observations and latent variables, depending on a parameter  $\theta \in \Theta$ .

The Expectation Maximization (EM) algorithm provides a numerical process to answer this problem by computing iteratively a sequence of estimates  $(\theta_n)_{n \in \mathbb{N}}$  which, under several conditions (see [12, 34]), converges towards the maximum likelihood estimate. It proceeds in two steps for each iteration  $k$ . First, in the Expectation step (E), the function

$$Q(\theta|\theta_{k-1}) = \int_Z \log(f(y, z, \theta)) p(y, z, \theta_{k-1}) dz$$

is computed where  $p$  is the conditional distribution of  $z$  given the observations:  $p(y, z, \theta) = f(y, z, \theta)/g(y, \theta)$ .  $\theta_k$  is then updated in the Maximization step (M) as the argument of the maximum of the function  $Q(\cdot|\theta_{k-1})$ .

The EM algorithm has been first introduced in [12]. Its properties have then been studied in numerous papers, see [6, 9, 23, 24, 27, 31, 34] among many other works.

In many cases, the (E) step is in fact intractable as we have no closed form for  $Q$ . Different algorithms, both deterministic and stochastic, have been introduced in the literature to overcome this problem. The Monte-Carlo EM ([33]) replaces the (E) step by computing a Monte Carlo approximation of  $Q$  using a large amount of simulated missing data  $z$ . Another possibility, more computationally efficient, is to use a Stochastic Approximation (SA) of the function  $Q$ . This SAEM algorithm has been introduced in [11] and the authors proved the convergence towards a local maximum of the incomplete likelihood with probability 1 under several hypotheses. It has later on been generalized in [15] in the case where we are not able to easily sample  $z$ . This new algorithm, called the SAEM Monte Carlo Markov Chain (SAEM-MCMC) replaces the sampling of  $z$  by one step of a Markov Chain targeting the conditional distribution  $p$ . Those two algorithms have then been applied in lots of different contexts: deformable models [3, 8, 10, 29], Independent Component Analysis [4] and in many medical problems (see [7, 14, 20, 30] among many others).

Among the hypotheses ensuring the convergence of most of these algorithms, and in particular our focus, the SAEM algorithm, one of the most restrictive is the necessity for the joint likelihood to belong to the curved exponential family. This writes:

$$f(y, z, \theta) = \exp(-\Psi(\theta) + \langle S(y, z), \Phi(\theta) \rangle), \quad (1.1)$$

where  $S$  is called a sufficient statistic of the model and  $\Phi$  and  $\Psi$  are two functions on  $\Theta$ . Similarly, the different extensions to the SAEM algorithm, and some of the EM algorithm, carry the same assumption [17, 18, 26, 28].

However, this hypothesis can in fact be a bottleneck in lots of situations. For example, it is not verified for heteroscedastic models [13, 16] nor with more complex models [8, 10, 21, 25, 29, 32]. Most of the authors then choose to compute the maximization step using a gradient descent. However, in that case, there is no theoretical guarantee of convergence. Moreover, the computational complexity increases. One needs to compute the gradient descent steps and compute the stochastic approximation of the complete likelihood while this function may not have a simple form. To solve this problem, [16] propose to transform the initial model to make it curved exponential.

Their solution consists in considering the parameter  $\theta$  as a realization of a Gaussian vector of mean  $\bar{\theta}$  and fixed variance  $\sigma^2$ .  $\theta$  then becomes an additional latent variable and the new parameter to estimate is  $\bar{\theta}$ . We call this new model the exponentialized model. It now belongs to the curved exponential family. However, as the likelihood of this exponentialized model is different, the function to maximize has also been modified. In particular, there is no guarantee that the new parameter to estimate is close to the initial one. Nevertheless, this trick has been successfully used in different situations ([1, 8, 10, 29] among others).

In this paper, we will study the maximum likelihood of this new exponentialized model and measure its distance to one of the maxima of the initial likelihood. More precisely, we will show that this distance goes to 0 as the variance  $\sigma^2$  of the exponentialized model tends to 0. We will also provide an upper bound to this error when  $\sigma$  is small enough. Finally, we will verify those results on an example. This example will show us

that a compromise must be done in the choice of  $\sigma$ . Indeed, if  $\sigma$  is too big, a substantial error is made in the estimation. However, for  $\sigma$  too small, despite the theoretical guarantees, the numerical convergence is difficult to obtain. To overcome this problem, we will present a new algorithm allowing a better estimation of the initial parameter  $\theta$  in a reasonable computation time.

## 2. PRESENTATION OF THE STOCHASTIC APPROXIMATION EXPECTATION MAXIMIZATION (SAEM) ALGORITHM

In this section, we recall the Stochastic Approximation Expectation Maximization (SAEM) algorithm, first presented in [11] and recall the hypotheses ensuring convergence. In the following, we suppose that the observation  $y$  belongs to  $\mathbb{R}^n$ , the latent variable  $z$  to  $\mathbb{R}^l$  and that the parameter space  $\Theta$  is an open subset of  $\mathbb{R}^p$  with  $n, l, p \in \mathbb{N}^*$ .

### 2.1. Expectation Maximization (EM) Algorithm

The original EM algorithm proposes to maximize a function defined *via*:

$$g(y, \theta) = \int_{\mathbb{R}^l} f(y, z, \theta) \mu(dz)$$

with  $f$  the joint likelihood of the model and  $\mu$  is a  $\sigma$ -finite measure on  $\mathbb{R}^l$ .

This situation is of interest to estimate the parameters of a statistical model using maximum likelihood estimates where the model depends on unobserved latent variables.

The Expectation-Maximization consists of iterations which guarantee an increase in  $g(\theta_k)$  at each step. Starting from  $\theta_0$ , the algorithm iterates:

- *Expectation.* Compute  $Q_k(\theta) = \int_{\mathbb{R}^l} \log(f(y, z, \theta)) p(y, z, \theta_k) dz$ .
- *Maximization.* Set  $\theta_{k+1} \in \operatorname{argmax} Q_k(\theta)$ .

where  $p$  is the conditional distribution of  $z$  given the observations:

$$p(y, z, \theta) = \begin{cases} f(y, z, \theta) / g(y, \theta) & \text{if } g(y, \theta) \neq 0 \\ 0 & \text{otherwise.} \end{cases}$$

### 2.2. SAEM algorithm

Because the expectation with respect to the conditional distribution  $p(y, z, \theta)$  is often intractable in practice, a different approach suggests replacing the E-step by a stochastic approximation on  $Q$ , starting from  $\theta_0$  and  $Q_0 = 0$ . This gives us the following algorithm:

- *Simulation.* Generate  $z_k$ , a realization of the hidden variable under the conditional density  $p(y, z, \theta_k)$ .
- *Approximation.* Update

$$Q_k(\theta) = Q_{k-1}(\theta) + \gamma_k (\log f(y, z_k, \theta) - Q_{k-1}(\theta)). \quad (2.1)$$

- *Maximization.* Set  $\theta_{k+1} \in \operatorname{argmax} Q_k(\theta)$ .

Convergence of this procedure is shown under the following hypotheses:

- (M1) The parameter space  $\Theta$  is an open subset of  $\mathbb{R}^p$ , and  $f$  can write:

$$f(y, z, \theta) = \exp(-\Psi(\theta) + \langle S(y, z), \Phi(\theta) \rangle), \quad (2.2)$$

where  $S(\cdot)$  is a Borel function taking its value in  $\mathcal{S}$ , an open subset of  $\mathbb{R}^{n_s}$ . In that case, we say that  $f$  belongs to the curved exponential family.

Moreover, the convex hull of  $S(\mathbb{R}^l)$  is included in  $\mathcal{S}$  and, for all  $\theta \in \Theta$ ,

$$\int_{\mathbb{R}^l} |S(y, z)| p(y, z, \theta) \mu(dz) < \infty.$$

(M2) The functions  $\Psi$  and  $\Phi$  are twice continuously differentiable on  $\Theta$ .

(M3) The function  $s : \Theta \rightarrow \mathcal{S}$  defined as:

$$s(\theta) = \int_{\mathbb{R}^l} S(y, z) p(y, z, \theta) \mu(dz)$$

is continuously differentiable on  $\Theta$ .

(M4) The observed log-likelihood  $l(\theta) := \log g(y, \theta)$  is continuously differentiable on  $\Theta$  and

$$\partial_\theta g(y, \theta) = \int_{\mathbb{R}^l} \partial_\theta f(y, z, \theta) \mu(dz).$$

(M5) There exists a function  $\hat{\theta} : \mathcal{S} \rightarrow \Theta$  such that  $\forall \theta \in \Theta, \forall s \in \mathcal{S}, L(s, \hat{\theta}(s)) \geq L(s, \theta)$ , with  $L(s, \theta) = -\Psi(\theta) + \langle s, \Phi(\theta) \rangle$ .

Moreover,  $\hat{\theta}$  is continuously differentiable on  $\mathcal{S}$ .

(SAEM1) For all  $k \geq 0$ ,  $0 \leq \gamma_k \leq 1$ ,  $\sum_{i=1}^{\infty} \gamma_k = \infty$  and  $\sum_{i=1}^{\infty} \gamma_k^2 < \infty$ .

(SAEM2)  $\hat{\theta} : \mathcal{S} \rightarrow \Theta$  and the observed-data log likelihood  $l : \theta \rightarrow \mathbb{R}$  are  $n_s$  times differentiable.

(SAEM3) For all positive Borel function  $\phi$ :

$$\mathbb{E}(\phi(z_{k+1}) | \mathcal{F}_k) = \int_{\mathbb{R}^l} \phi(z) p(z, \theta_k) \mu(dz),$$

where  $z_k$  is the missing value simulated at step  $k$  under the conditional density  $p(y, z, \theta_{k-1})$  and  $\mathcal{F}_n$  is the family of  $\sigma$ -algebra generated by the random variables  $S_0, z_1, \dots, z_n$ .

(SAEM4) For all  $\theta \in \Theta$ ,  $\int_{\mathbb{R}^l} \|S(y, z)\|^2 p(y, z, \theta) \mu(dz) < \infty$  and  $\Gamma(\theta) := \text{Cov}_\theta(S(y, z))$  is continuous with respect to  $\theta$ .

With the hypothesis (M1) specifying the form of the complete likelihood and (M5) giving us the existence of a maximizer  $\hat{\theta}$ , the algorithm can take a simpler form. Indeed, using the fact that  $Q$  is fully defined by a sufficient statistic  $S$ , we remark, by linearity, that the stochastic approximation (2.1) is only applied on this sufficient statistic. Similarly, the maximization step can be rewritten using only the sufficient statistic and  $\hat{\theta}$ . This gives the following algorithm:

- *Simulation.* Generate  $z_k$ , a realization of the hidden variable under the conditional density  $p(y, z, \theta_k)$ .
- *Approximation.* Update  $S_k = S_{k-1} + \gamma_k(S(y, z) - S_{k-1})$
- *Maximization.* Set  $\theta_{k+1} = \hat{\theta}(S_k)$ .

We finally assume the following hypothesis:

- (A) With probability 1,  $\text{clos}((S_k)_{k \geq 1})$  is a compact subset of  $\mathcal{S}$ .

**Remark 2.1.** The assumption (A) can be relaxed by projecting the sequence  $(S_k)_{k \in \mathbb{N}}$  on increasing compacts. See [5] for more details.

Under the hypotheses (M1)–(M5), (SAEM1)–(SAEM4) and (A), it was shown in [5] that the distance between the sequence generated by the SAEM and the set of stationary point of the observed likelihood  $g$  converges almost surely towards 0.

However, in numerous cases, even quite simple [13, 16], the joint likelihood  $f$  does not verify the hypothesis (M1) as it does not belong to the curved exponential family. In the next Section, we will present a trick allowing us to approximate the maximum likelihood when (M1) is not verified.

In the following, to simplify the notations, we no longer write the variable  $y$  in the different expressions.

### 2.3. Exponentialization process

We now denote by  $(\theta, \psi)$  the parameters of  $g$  where  $\theta \in \Theta$ ,  $\psi \in \Omega = \mathbb{R}^m$ , and we tackle the case where the model cannot be written under the curved exponential form (2.3) because of the parameter  $\psi$ . In that case, the log-likelihood can only be written as:

$$f(z, \theta, \psi) = \exp(-\Psi(\theta) + \langle S(z), \Phi(\theta) \rangle) h(z, \psi) \quad (2.3)$$

and  $f$  does not belong to the curved exponential family.

Here, some parameters  $\theta$  are separable from the latent variables  $z$  and do not require further transformation. Other variables  $\psi$  are at the source of the computational problem and the exponentialization process will only apply on those parameters. It must be noticed that, in some cases,  $\theta$  can be empty.

The trick proposed in [16] is to consider  $\psi$  as a Gaussian random variable  $\psi \sim \otimes \mathcal{N}(\bar{\psi}, \sigma^2)$ , where the notation  $\otimes \mathcal{N}(., .)$  denotes a multivariate Gaussian distribution with diagonal covariance matrix. Hence, in this augmented model,  $\psi$  is no longer a parameter but becomes an additional latent variable while a new parameter  $\bar{\psi}$  appears.

The resulting perturbed statistical model is curved exponential, with augmented parameters  $\hat{\theta} = (\theta, \bar{\psi})$  and augmented random latent variables  $\hat{z} = (z, \psi)$ .

The variance  $\sigma^2$  is chosen by the user, and should be reasonably small in order to minimally perturb the original model. In practice, this variance should at the same time be chosen reasonably large in order to speed-up the parameter estimation (see experiments in Sect. 4).

The complete log-likelihood of this exponentialized model then writes:

$$\log f_\sigma(\hat{z}, \hat{\theta}) = -\Psi(\theta) + \langle S(z), \Phi(\theta) \rangle + \log(h(z, \psi)) - \frac{\|\psi - \bar{\psi}\|^2}{2\sigma^2}. \quad (2.4)$$

It is easy to check that the complete log-likelihood now belongs to the curved exponential family with sufficient statistics:  $(S(z), \psi)$ . Concerning the parameter  $\theta$ , the maximization is done as usual:  $\theta_{k+1} = \hat{\theta}(S_k)$  with  $S_k$  the stochastic approximation of the  $(S(z_i))_{i \leq k}$ . The update of the parameter  $\bar{\psi}$  can for its part be written as:

$$\bar{\psi}_{k+1} = \bar{\psi}_k + \gamma_k(\psi_k - \bar{\psi}_k). \quad (2.5)$$

If we suppose that this augmented model satisfies the hypotheses (M1)–(M5), (SAEM1)–(SAEM4) and (A), we know, using the theorem proved in [5], that it will converge towards a critical point of its incomplete likelihood. However, if this process is used in several applications [19], there is in fact no guarantee that the algorithm will converge towards a critical point of the incomplete log-likelihood of the initial model.

In the following section, we show that, in general, the parameter returned by the SAEM on the exponentialized model is indeed not a maximum likelihood of the initial model. However, when  $\sigma$  goes to 0, it converges towards

a critical point of the incomplete log likelihood of the initial model. We also give an upper bound of the error made by this process for  $\sigma$  small.

It is interesting to notice that, even if this proof is done in the context of the SAEM algorithm, the same results can be obtained for the MCMC-SAEM [15] as well as for the Approximate SAEM [2].

### 3. DISTANCE BETWEEN THE LIMIT POINT AND THE NEAREST CRITICAL POINT

In this section, we first present an equation satisfied by the limit of the sequence of estimated parameters of the SAEM algorithm for the exponentialized model. Using this equation, we will then give an upper bound on the distance between this limit point and the nearest critical point of the incomplete likelihood of the non-exponential model. This upper bound will in particular show us that this distance tends to 0 when  $\sigma$  goes to 0.

#### 3.1. Equation verified by the limit

We now state a theorem giving us an equation satisfied by the limit parameter estimated by the SAEM algorithm applied on the exponentialized model. It is important to remark that, if the set of the critical points of  $l$  is finite then, the SAEM algorithm converges almost surely towards one of them (and not only towards a point at zero distance). Hence, we can study the parameters returned by the SAEM on the exponential model:  $\bar{\psi}_\sigma$  and look at their behaviour when  $\sigma$  goes to 0.

**Theorem 3.1.** *Assume that the exponentialized model with variance  $\sigma$  verifies the hypotheses (M1)–(M5), (SAEM1)–(SAEM4), (A) and that  $\Omega = \mathbb{R}^m$ . Assume also that, for all  $\sigma > 0$ ,*

$$\mathcal{L}_\sigma := \{(\theta, \bar{\psi}) \in \Theta \times \Omega \mid \partial_{\theta, \bar{\psi}} l_\sigma(\theta, \bar{\psi}) = 0\}$$

is finite where  $l_\sigma$  refers to the observed log-likelihood of the exponentialized model of variance  $\sigma$ . Then, the sequence returned by the SAEM algorithm converges almost surely towards  $(\theta_\infty, \bar{\psi}_\sigma)$ , solutions of the following set of equations:  $\forall 1 \leq k \leq m$ ,

$$\int_{\mathbb{R}^m} v_k g(\theta_\infty, \bar{\psi}_\sigma + v) \exp\left(-\frac{\|v\|^2}{2\sigma^2}\right) dv = 0, \quad (3.1)$$

where  $v_k$  is the  $k$ -th coordinate of  $v \in \mathbb{R}^m$  and  $g(\theta, \psi) = \int_Z f(z, \theta, \psi) \mu(dz)$ .

**Remark 3.2.** Here, we suppose  $\Omega = \mathbb{R}^m$  to be able to define a Gaussian distribution on  $\Omega$ . The following proofs would be adaptable as long as one can define such a gaussian distribution, necessary for applying the exponentialization trick.

*Proof.* The update of  $S_k, \psi_k$  can easily be seen as a Robbins Monro update:

$$\begin{cases} \bar{\psi}_{k+1} = \bar{\psi}_k + \gamma_k v_k \\ S_{k+1} = S_k + \gamma_k (S(z_k) - S_{k-1}) \end{cases}$$

where  $z_k, v_k$  are sampled following the conditional law  $f_\sigma(z, \psi + \bar{\psi}_k | \theta_k, \bar{\psi}_k)$ .

Under the hypotheses explained Section 2.2, it has been shown that this Robbin Monroe approximation verifies  $\lim_{k \rightarrow \infty} d((\theta_k, \bar{\psi}_k), \mathcal{L}_\sigma) = 0$ . Moreover, because  $\mathcal{L}_\sigma$  is finite, [11] show that the sequence  $(\theta_k, \bar{\psi}_k)_{k \geq 0}$  converges almost surely towards a point  $(\theta_\infty, \bar{\psi}_\sigma) \in \Theta \times \Omega$  (Thm. 6). Using the regularity of  $l_\sigma$ , we deduce that those parameters verify  $\partial_{\theta, \bar{\psi}} l_\sigma(\theta_\infty, \bar{\psi}_\sigma) = 0$ .

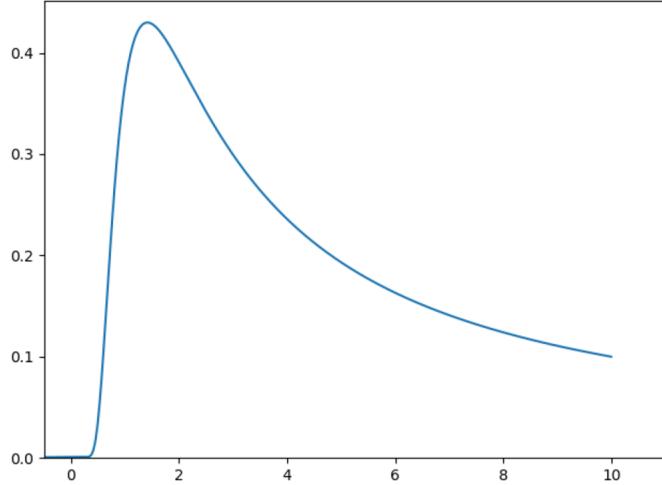


FIGURE 1. Function  $g$  studied subsection 3.2, with a maximum reached at  $\sqrt{2}$ .

By replacing  $l_\sigma$  by its value in this equation and using the assumption (M4), we find, for all  $1 \leq k \leq m$ :

$$\int_{\mathbb{R}^m} (v_k - \bar{\psi}_\sigma) g(z, \theta_\infty, v) \exp\left(-\frac{\|v - \bar{\psi}_\sigma\|^2}{2\sigma^2}\right) dv dz = 0$$

Using a change of variable, we finally find the expected result.  $\square$

**Proposition 3.3.** *Suppose that a point  $\psi \in \mathbb{R}^m$  verifies:*

$$\forall v, \theta \in \mathbb{R}^m \times \Theta, g(\theta, \psi + v) = g(\theta, \psi - v). \quad (3.2)$$

*Then,  $\psi$  is solution to equation (3.1) and can be the parameter returned by the exponential model.*

**Remark 3.4.** It is not necessarily the only possibility of returned parameter. Several values of  $\psi$  could be solutions of equation (3.1).

In particular, a solution of (3.2) is a critical point of  $g$ . However, a critical point of  $g$  is not always solution of such an equation and will not always be a solution of equation (3.1). It is in fact easy to find cases where the maximum is not a solution to equation (3.1) and hence where the exponentialized model introduces a bias in the estimation of maximum likelihood.

We introduce next subsection a function  $g$  presenting such a behaviour and we explain the heuristics behind Theorem 3.1.

### 3.2. Heuristics

We want to compare the solution of equation (3.1) to a maximum of the function  $g$ . Because of the form of  $f$  supposed in equation (2.3), we see that we can maximize  $g$  in  $\theta$  and  $\psi$  independently of the other. In particular, we still immediately have  $\theta_\infty \in \operatorname{argmax}_\theta g(\theta, \psi)$  (independent of  $\psi$  as can be seen in Eq. (2.3)).

To explain equation (3.1), we introduce the function  $g : v \mapsto \frac{1}{v} \exp(-\frac{1}{v^2})$  presented Figure 1. This function has a maximum for  $\psi = \sqrt{2}$  but is not symmetric around it. We will see that this lack of symmetry around its

maximum causes  $\sqrt{2}$  not to be solution of equation (3.1). We will look at the value of the integral:

$$\int_{\Omega} vg(\bar{\psi} + v) \exp\left(-\frac{v^2}{2\sigma^2}\right) dv$$

for different values of  $\bar{\psi}$  and  $\sigma$ .  $\bar{\psi}$  is a solution of equation (3.1) if and only if this integral is null. It is interesting to remark that one can consider this integral as an expectation if normalized. On the different figures of this section, we add a green line at  $x = 0$  to visualize the value of this integral.

First we look at the case  $\bar{\psi} = \sqrt{2}$ , the argmax of  $g$ , and  $\sigma = 1$  on Figure 2a. In that case, because  $g$  is not symmetric around its maximum,  $v \mapsto g(\sqrt{2} + v) \exp\left(-\frac{v^2}{2\sigma^2}\right)$  is not symmetric either. In particular, it means that  $\sqrt{2}$  is not a solution of equation (3.1) as the integral is strictly positive. This lack of symmetry will stay true for any value of  $\sigma$ , even small, causing  $\sqrt{2}$  never to be solution of equation (3.1).

We now interest ourselves in the case where  $\bar{\psi}$  is not the argmax of  $g$  by taking  $\bar{\psi} = 1$  and  $\sigma = 1$  on Figure 2b. Because  $g$  is strictly increasing at 1,  $v \mapsto g(1 + v) \exp\left(-\frac{v^2}{2\sigma^2}\right)$  increases at 0. As it decreases slower than it increases, the integral is strictly positive and 1 is not a solution of equation (3.1). The same behaviour would be observed for any point before  $\sqrt{2}$  and any value of  $\sigma$ .

We now look at a value bigger than the maximum:  $\bar{\psi} = 4$  and  $\sigma = 3.67$  on Figure 2c. This time,  $v \mapsto g(4 + v) \exp\left(-\frac{v^2}{2\sigma^2}\right)$  decreases at 0. But  $g$  decreases slower than it increases. Hence, because we have taken  $\bar{\psi} > \sqrt{2}$ , this difference of variation is compensated and, for this particular value of  $\sigma$ , 4 is solution of equation (3.1).

Finally, let us take a smaller value of  $\sigma$  as in Figure 2d. This time, the integral is negative. Indeed, the difference of variation before and after the maximum is now way smaller and does not compensate the decrease of  $g$  at 0. To have a solution of equation (3.1) for this particular value of  $\sigma$ , we would need to choose a value of  $\bar{\psi}$  smaller. This suggests that, as  $\sigma$  goes to 0, the solution of equation (3.1) is closer to  $\sqrt{2}$ .

From these examples, we can deduce two things. First, the argmax of  $g$  is not always solution of the equation (3.1) even for small values of  $\sigma$  when there is a difference in the speed of variation before and after this maximum. Moreover, when  $\sigma$  goes to 0, it seems that a parameter closer and closer to the argmax is solution of equation (3.1).

We illustrate this behaviour by plotting the exact value of the solution of equation (3.1) as a function of  $\sigma$  in Figure 3 for this particular function  $g$ .

In the following, we write  $\psi_M$  the critical point of  $g(\theta, \psi)$  minimizing the distance to  $\bar{\psi}_\sigma$ .

Using the heuristics presented above, we will, in the next Section, state the theorem giving us an upper bound on the distance to the nearest critical point of  $g$ . We will then prove it in the case  $\Omega = \mathbb{R}$ . A more general proof in  $\mathbb{R}^m$  for  $m \geq 2$  is given in the annex.

### 3.3. Upper bound on the distance between $\bar{\psi}_\sigma$ and the nearest critical point of $g$

**Theorem 3.5.** – Assume that the exponential model verifies the hypotheses (M1)-(M5), (SAEM1)-(SAEM4) and (A). Assume also that, for all  $\sigma > 0$ ,  $\mathcal{L}_\sigma$  is finite, that  $\mathcal{L} := \{\psi \in \mathbb{R}^m \mid \partial_\psi g(\theta, \psi) = 0\}$  is compact and that there exists  $K$  compact such that,  $\forall \sigma > 0$ ,  $\bar{\psi}_\sigma \in K$ . Then,

$$d(\bar{\psi}_\sigma, \mathcal{L}) \xrightarrow[\sigma \rightarrow 0]{} 0.$$

– Assume also that  $\mathcal{L}$  is finite and that, for all  $\psi_M \in \mathcal{L}$ , there exists an integer  $l_M$  such that  $g$  is  $l_M$ -times continuously differentiable and such that

$$\forall k \leq m, \exists i \leq l_M \text{ with } \frac{\partial^i g}{\partial \psi_k^i}(\psi_M) \neq 0.$$

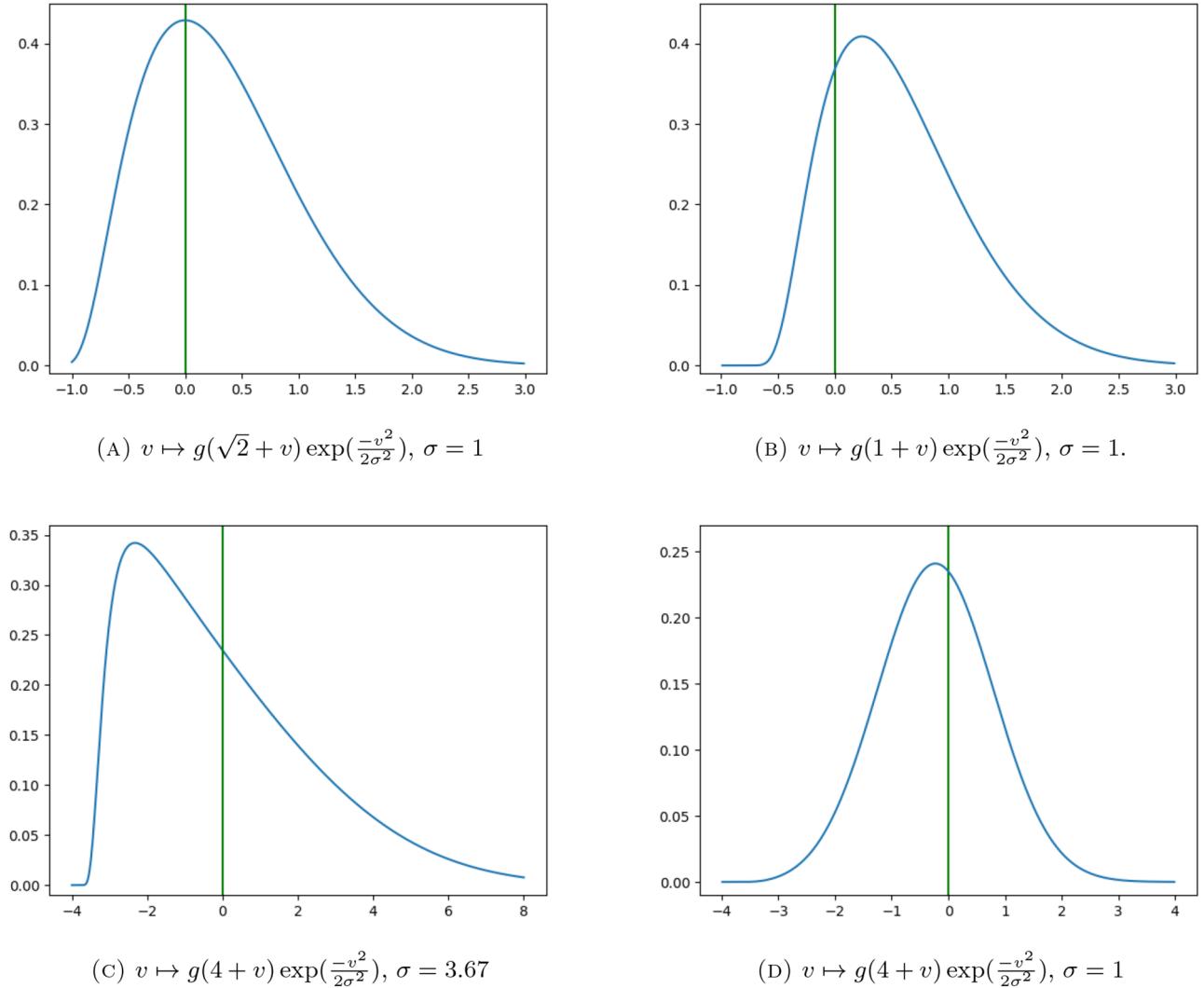


FIGURE 2. Plot of  $v \mapsto g(\bar{\psi} + v) \exp\left(\frac{-v^2}{2\sigma^2}\right)$  for different values of  $\bar{\psi}$  and  $\sigma$ . The green line highlights the point  $x = 0$ . On the first line, we can see that  $\sqrt{2}$  is not a solution of equation (3.1) because  $g$  increases quicker than it decreases. On the second line, since  $g$  is increasing at 1, we see that 1 is not solution of (3.1). Finally, on subfigure 2c,  $\bar{\psi} = 4$  is solution of equation (3.1) with  $\sigma = 3.67$ . On subfigure 2d, for a smaller value of  $\sigma$ , the integral (3.1) is this time negative.

We write  $l = \max_{\psi_M \in \mathcal{L}} l_M$ .

Then, there exists  $c > 0$  such that, for  $\sigma$  small enough,

$$d(\bar{\psi}_\sigma, \mathcal{L}) \leq c\sigma^{\frac{2}{l+2}}. \quad (3.3)$$

– Suppose that  $v \mapsto g(\theta_\infty, v)$  and  $v \mapsto v_k g(\theta_\infty, v)$  are integrable for all  $k$  between 1 and  $m$ . Then, we have the following approximation of  $\bar{\psi}_\sigma$  when  $\sigma$  goes to infinity: for all  $1 \leq k \leq m$ ,

$$(\bar{\psi}_\sigma)_k \xrightarrow[\sigma \rightarrow \infty]{} \frac{\int_{\mathbb{R}^m} v_k g(\theta_\infty, v) dv}{\int_{\mathbb{R}^m} g(\theta_\infty, v) dv}.$$

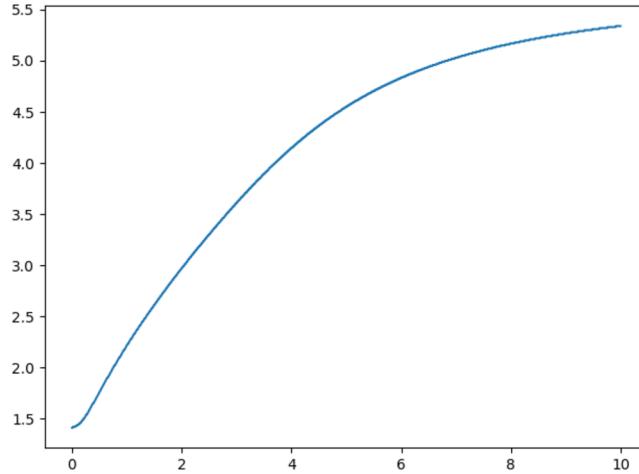


FIGURE 3. Solution of equation (3.1) as a function of  $\sigma$  for the function  $g$  studied subsection 3.2.

**Remark 3.6.** When  $m = 1$ ,  $l_M$  is the smallest integer such that the  $l_M$ -th derivative of  $g(\theta_\infty, \cdot)$  at  $\psi_M$  is not 0. The inequality (3.3) indicates that the convergence will be slower when the function to maximize behaves as a flat curve around the maximum, which was expected.

If such a  $l_M$  does not exist, it means that  $g$  is constant in at least one direction around  $\psi_M$ .

**Remark 3.7.** We need to take a maximum over all  $\psi_M$  in  $\mathcal{L}$  since, from one  $\sigma$  to another,  $\psi_\sigma$  can approach a different maximum  $\psi_M \in \mathcal{L}$ . It is also why the upper bound depends on this maximum  $l$ . It is constrained by the critical point for which the convergence is the slowest.

**Remark 3.8.** For  $m = 1$ , we have the exact value of the constant in (3.3):

$$c = \max_{\psi_M \in \mathcal{L}} \left[ \left( 12(l_M - 1)! \frac{\|g\|_\infty}{|\partial_\psi^{l_M} g(\theta_\infty, \psi_M)|} \right)^{\frac{1}{l_M+2}} \right].$$

**Remark 3.9.** The results presented here are still true in the case of the SAEM-MCMC algorithm. Indeed, in that case, the limit parameters still verify  $\lim_{k \rightarrow \infty} d((\theta_k, \bar{\psi}_k), \mathcal{L}_\sigma) = 0$ . Hence, the equation (3.1) and the Theorem 3.5 can still be verified using the same steps as for the SAEM algorithm.

In the following, we will present the proof in the case  $m = 1$ . The proof in the multi-dimensional case follows the same ideas than in dimension one but is more technical. It is presented in the annex.

*Proof.* We present the proof in the case  $m = 1$ . As the maximum does not depend of  $\theta_\infty$  and to simplify notations, we will forget the variable  $\theta$  in  $g$  and use  $g(\psi) = g(\theta_\infty, \psi)$ .

**First step:**  $d(\bar{\psi}_\sigma, \mathcal{L}) \xrightarrow[\sigma \rightarrow 0]{} 0$

We suppose that  $\bar{\psi}_\sigma$  is never a critical point of  $g$  for  $\sigma$  small enough. Otherwise, we directly have the result. The equation (3.1) writes:

$$\int_{\mathbb{R}} v g(\bar{\psi}_\sigma + v) \exp\left(-\frac{v^2}{2\sigma^2}\right) dv = 0.$$

The first step is to show that  $d(\bar{\psi}_\sigma, \mathcal{L}) \xrightarrow[\sigma \rightarrow 0]{} 0$ . By contradiction, even if it means extracting a subsequence, we can suppose that there exists  $c > 0$  such that  $\forall \sigma > 0$ ,  $d(\bar{\psi}_\sigma, \mathcal{L}) > 3c$ .

Because there is no critical point between  $\bar{\psi}_\sigma - c$  and  $\bar{\psi}_\sigma + c$ ,  $g$  is either increasing or decreasing on  $[\bar{\psi}_\sigma - c, \bar{\psi}_\sigma + c]$ . We first suppose it is increasing. In particular,  $K_0 := K \setminus \{y \mid d(y, \mathcal{L}) < c\}$  is compact and thus  $c_0 := \inf\{g'(y) \mid y \in K_0, g'(y) \geq 0\} > 0$ . According to equation (3.1), the integral on  $[-c, c]$  must have the same absolute value as the integral on  $[-c, c]^c$ . However, we will show that, when  $\sigma$  goes to zero, the first one converges towards 0 much more slowly than the second one. Indeed,

$$\begin{aligned} \int_{|v| \geq c} v g(\bar{\psi}_\sigma + v) \exp\left(-\frac{v^2}{2\sigma^2}\right) dv &\geq \int_{v \leq -c} v g(\bar{\psi}_\sigma + v) \exp\left(-\frac{v^2}{2\sigma^2}\right) dv \\ &\geq \|g\|_\infty \int_{v \leq -c} v \exp\left(-\frac{v^2}{2\sigma^2}\right) dv \\ &\geq -\sigma^2 \|g\|_\infty \exp\left(-\frac{c^2}{2\sigma^2}\right). \end{aligned}$$

On the other hand, we have:

$$\int_{|v| \leq c} v g(\bar{\psi}_\sigma + v) \exp\left(-\frac{v^2}{2\sigma^2}\right) dv = \int_{0 \leq v \leq c} v (g(\bar{\psi}_\sigma + v) - g(\bar{\psi}_\sigma - v)) \exp\left(-\frac{v^2}{2\sigma^2}\right) dv.$$

Using the mean value theorem, for all  $0 \leq v \leq c$ , there exists  $\tilde{\psi}_v \in [\bar{\psi}_\sigma - v, \bar{\psi}_\sigma + v] \subset K_0$  such that  $g(\bar{\psi}_\sigma + v) - g(\bar{\psi}_\sigma - v) = 2v g'(\tilde{\psi}_v) \geq 2c_0 v$ . Hence, we find:

$$\int_{|v| \leq c} v g(\bar{\psi}_\sigma + v) \exp\left(-\frac{v^2}{2\sigma^2}\right) dv \geq 2c_0 \int_{0 \leq v \leq c} v^2 \exp\left(-\frac{v^2}{2\sigma^2}\right) dv.$$

But, using an integration per part and defining

$$\text{erf}(x) := \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt,$$

we have:

$$\int_{|v| \leq c} v g(\bar{\psi}_\sigma + v) \exp\left(-\frac{v^2}{2\sigma^2}\right) dv \geq 2c_0 \sigma^2 \left[ -c \exp\left(-\frac{c^2}{2\sigma^2}\right) + \sigma \frac{\sqrt{\pi}}{2} \text{erf}\left(\frac{c}{\sqrt{2}\sigma}\right) \right].$$

Hence, because

$$\int_{|v| \leq c} v g(\bar{\psi}_\sigma + v) \exp\left(-\frac{v^2}{2\sigma^2}\right) dv = - \int_{|v| \geq c} v g(\bar{\psi}_\sigma + v) \exp\left(-\frac{v^2}{2\sigma^2}\right) dv,$$

we have:

$$\|g\|_\infty \exp\left(-\frac{c^2}{2\sigma^2}\right) \geq 2c_0 \left[ -c \exp\left(-\frac{c^2}{2\sigma^2}\right) + \sigma \frac{\sqrt{\pi}}{2} \text{erf}\left(\frac{c}{\sqrt{2}\sigma}\right) \right].$$

It is easy to find the same inequality if  $g$  is decreasing on  $[\bar{\psi}_\sigma - c, \bar{\psi}_\sigma + c]$ . Indeed, in that case, using the fact that the integral on  $\{v \leq c\}$  is negative and using the same inequalities as above for the integral on  $\{v \geq c\}$ , we first show that

$$\int_{|v| \geq c} v g(\bar{\psi}_\sigma + v) \exp\left(-\frac{v^2}{2\sigma^2}\right) dv \leq \sigma^2 \|g\|_\infty \exp\left(-\frac{c^2}{2\sigma^2}\right).$$

Then, by considering this time  $c_1 := \sup\{g'(y) \mid y \in K_0, g'(y) \leq 0\} < 0$ , and using again the mean value theorem, we find:

$$\begin{aligned} \int_{|v| \leq c} vg(\bar{\psi}_\sigma + v) \exp\left(-\frac{v^2}{2\sigma^2}\right) dv &\leq 2c_1 \int_{0 \leq v \leq c} v^2 \exp\left(-\frac{v^2}{2\sigma^2}\right) dv \\ &\leq 2c_1 \sigma^2 \left[ -c \exp\left(-\frac{c^2}{2\sigma^2}\right) + \sigma \frac{\sqrt{\pi}}{2} \operatorname{erf}\left(\frac{c}{\sqrt{2}\sigma}\right) \right]. \end{aligned}$$

Hence, for all  $\sigma > 0$ , there exists  $C := 2 \max(c_0, -c_1) > 0$  such that:

$$\frac{\|g\|_\infty}{\sigma} \exp\left(-\frac{c^2}{2\sigma^2}\right) \geq C \left[ -\frac{c}{\sigma} \exp\left(-\frac{c^2}{2\sigma^2}\right) + \frac{\sqrt{\pi}}{2} \operatorname{erf}\left(\frac{c}{\sqrt{2}\sigma}\right) \right].$$

By taking  $\sigma$  to 0 and using the fact that  $\operatorname{erf}(x) \xrightarrow{x \rightarrow \infty} 1$ , we find  $C \leq 0$  which is a contradiction. Hence, we have proved that

$$d(\bar{\psi}_\sigma, \mathcal{L}) \xrightarrow{\sigma \rightarrow 0} 0.$$

The next step is to find an upper bound on  $d(\bar{\psi}_\sigma, \mathcal{L})$ .

### Second step: Search of the upper bound

In the following, we will suppose that the critical point towards which  $\bar{\psi}_\sigma$  converges is a maximum. In practice, it will always be the case as any other critical point would be unstable numerically. Theoretically, a set of conditions (LOC1)-(LOC3) are given in [11] insuring the convergence towards a local maximum.

We write  $\psi_M$  the closest critical point to  $\bar{\psi}_\sigma$  and  $\alpha_\sigma = |\bar{\psi}_\sigma - \psi_M|$ . We also write  $l_M$  the smallest integer such that  $g^{(l_M)}(\psi_M) \neq 0$ . Moreover, as explained above, we assume that  $\psi_M$  is a maximum. It must be remarked that  $\psi_M$  depends on  $\sigma$ . However, as  $\mathcal{L}$  is finite, we will be able to consider maxima at the end of the proof. Since we assume  $\psi_M$  maximum,  $l_M$  is even and, for  $\sigma$  small enough, since  $g^{(l_M)}$  is continuous,  $\forall v \in [\bar{\psi}_\sigma - \alpha_\sigma, \bar{\psi}_\sigma + \alpha_\sigma]$ ,

$$g^{(l_M)}(v) \leq \frac{1}{2} g^{(l_M)}(\psi_M) := -c_M < 0.$$

As before, we will split up the integral (3.1) in two parts:  $\{v \mid |v| < \alpha_\sigma\}$  and  $\{v \mid |v| > \alpha_\sigma\}$ . The idea behind the computations is that  $\alpha_\sigma$  cannot be too big without making the absolute value of the integral on  $\{v \mid |v| < \alpha_\sigma\}$  strictly superior than the one on  $\{v \mid |v| > \alpha_\sigma\}$ .

On  $\{v \mid |v| > \alpha_\sigma\}$  we can use the same upper and lower bounds as before to find:

$$\left| \int_{|v| \geq \alpha_\sigma} vg(\bar{\psi}_\sigma + v) \exp\left(-\frac{v^2}{2\sigma^2}\right) dv \right| \leq \sigma^2 \|g\|_\infty \exp\left(-\frac{\alpha_\sigma^2}{2\sigma^2}\right). \quad (3.4)$$

On  $\{v \mid |v| < \alpha_\sigma\}$ , we use twice the mean value theorem to find, for any  $v \in [0, \alpha_\sigma]$ , there exist  $\tilde{\psi}_v^0 \in [\bar{\psi}_\sigma - v, \bar{\psi}_\sigma + v]$  and  $\tilde{\psi}_v^1 \in [\bar{\psi}_\sigma - \alpha_\sigma, \bar{\psi}_\sigma + \alpha_\sigma]$  such that:

$$\begin{aligned} g(\bar{\psi}_\sigma + v) - g(\bar{\psi}_\sigma - v) &= 2vg'(\tilde{\psi}_v^0) = 2v(g'(\tilde{\psi}_v^0) - g'(\psi_M)) \\ &= 2v(\tilde{\psi}_v^0 - \psi_M)^{l_M-1} g^{(l_M)}(\tilde{\psi}_v^0) / (l_M - 1)! \\ &\geq 2v(\bar{\psi}_\sigma + v - \psi_M)^{l_M-1} g^{(l_M)}(\tilde{\psi}_v^1) / (l_M - 1)!. \end{aligned}$$

We first suppose that  $g$  is increasing on  $[\bar{\psi}_\sigma - \alpha_\sigma, \bar{\psi}_\sigma + \alpha_\sigma]$ . Then,  $\alpha_\sigma = \psi_M - \bar{\psi}_\sigma$  and:

$$g(\bar{\psi}_\sigma + v) - g(\bar{\psi}_\sigma - v) \geq \frac{2c_M}{(l_M - 1)!} v (\alpha_\sigma - v)^{l_M - 1}.$$

Hence, computing the integral (3.1) on  $\{v \mid |v| < \alpha_\sigma\}$ , we find:

$$\begin{aligned} \int_{|v| \leq \alpha_\sigma} v g(\bar{\psi}_\sigma + v) \exp\left(-\frac{v^2}{2\sigma^2}\right) dv &\geq \frac{2c_M}{(l_M - 1)!} \int_0^{\alpha_\sigma} v^2 (\alpha_\sigma - v)^{l_M - 1} \exp\left(-\frac{v^2}{2\sigma^2}\right) dv \\ &\geq \frac{2c_M}{(l_M - 1)!} \alpha_\sigma^{l_M + 2} \int_0^1 v^2 (1 - v)^{l_M - 1} \exp\left(-\frac{\alpha_\sigma^2 v^2}{2\sigma^2}\right) dv \\ &\geq \frac{2c_M}{(l_M - 1)!} \alpha_\sigma^{l_M + 2} \exp\left(-\frac{\alpha_\sigma^2}{2\sigma^2}\right) \int_0^1 v^2 (1 - v) dv. \end{aligned}$$

Finally, by combining this inequality and (3.4), we find:

$$\sigma^2 \|g\|_\infty \exp\left(-\frac{\alpha_\sigma^2}{2\sigma^2}\right) \geq \frac{c_M}{6(l_M - 1)!} \alpha_\sigma^{l_M + 2} \exp\left(-\frac{\alpha_\sigma^2}{2\sigma^2}\right).$$

Hence, if  $\sigma \leq 1$ ,

$$\begin{aligned} \alpha_\sigma &\leq \left( 6(l_M - 1)! \frac{\|g\|_\infty}{c_M} \right)^{1/(l_M + 2)} \sigma^{\frac{2}{l_M + 2}} \\ &\leq \max_{\psi_M \in \mathcal{L}} \left( \left( 6(l_M - 1)! \frac{\|g\|_\infty}{c_M} \right)^{1/(l_M + 2)} \right) \sigma^{\frac{2}{l+2}}. \end{aligned}$$

Because  $\mathcal{L}$  is finite, we indeed have a maximum which is strictly positive.

In the case where  $g$  is decreasing on  $[\bar{\psi}_\sigma - \alpha_\sigma, \bar{\psi}_\sigma + \alpha_\sigma]$ , we have  $\alpha_\sigma = \bar{\psi}_\sigma - \psi_M$  and it is easy to show that we have this time

$$g(\bar{\psi}_\sigma + v) - g(\bar{\psi}_\sigma - v) \leq -2c_M v (\alpha_\sigma - v)^{l_M - 1} / (l_M - 1)!.$$

Hence, we can use the same inequalities as before to find again:

$$\alpha_\sigma \leq \max_{\psi_M \in \mathcal{L}} \left( \left( 6(l_M - 1)! \frac{\|g\|_\infty}{c_M} \right)^{1/(l_M + 2)} \right) \sigma^{\frac{2}{l+2}}.$$

### Third step: Approximation when $\sigma$ goes to infinity

We use again the equation (3.1). For all  $\sigma > 0$ ,

$$\int_{\mathbb{R}} v g(\bar{\psi}_\sigma + v) \exp\left(-\frac{v^2}{2\sigma^2}\right) dv = 0.$$

Using the change of variable  $\bar{\psi}_\sigma + v$ , we find:

$$\bar{\psi}_\sigma = \frac{\int_{\mathbb{R}} v g(v) \exp\left(-\frac{(v - \bar{\psi}_\sigma)^2}{2\sigma^2}\right) dv}{\int_{\mathbb{R}} g(v) \exp\left(-\frac{(v - \bar{\psi}_\sigma)^2}{2\sigma^2}\right) dv}.$$

But  $\bar{\psi}_\sigma$  is supposed to stay in a compact so,  $\forall v \in \mathbb{R}$ ,  $\exp\left(-\frac{(v-\bar{\psi}_\sigma)^2}{2\sigma^2}\right) \xrightarrow[\sigma \rightarrow \infty]{} 1$ . Using the integrability of  $g$  and  $v \mapsto vg(v)$ , it is easy to conclude using the dominated convergence theorem.  $\square$

#### 4. SIMULATION OF A COUNTER EXAMPLE

In this Section, we demonstrate that the maximum likelihood of  $g$  is indeed not reached by the SAEM algorithm on the exponentialized model on a concrete situation.

##### 4.1. Application of the SAEM algorithm to the exponentialized model

We choose to study a heteroscedastic model where the variance depends on the observation. This model has been used in [16] in order to analyze the growth of orange trees. The parameters to estimate are the age  $\beta_1$  at half asymptotic trunk circumference  $\psi_i$  and the grow scale  $\beta_2$  of  $n$  orange trees according to the measurement of their circumference  $y_{i,j}$  at  $m$  different ages  $x_j$ .

We suppose that our observation  $y_{i,j}$  verifies, for  $i$  between 1 and  $n$  and  $j$  between 1 and  $m$ :

$$y_{i,j} = \frac{\phi_i}{1 + \exp\left(-\frac{x_j - \beta_1}{\beta_2}\right)} (1 + \varepsilon_{i,j}),$$

where  $\varepsilon_{i,j}$  are independent noises of distribution  $\mathcal{N}(0, \sigma_\varepsilon^2)$  of variance  $\sigma_\varepsilon^2$  supposed to be known.  $\phi_i$  is treated as a random effect and is supposed to follow a Gaussian distribution of mean  $\mu$  to estimate and known variance  $\tau^2$ .

Such a model cannot be written in an exponential form due to the parameters  $\beta_1$  and  $\beta_2$  and we will hence consider an exponentialized model where  $\beta_1$  and  $\beta_2$  are considered as random effects with  $\beta_1 \sim \mathcal{N}(\bar{\beta}_1, \sigma^2)$  and  $\beta_2 \sim \mathcal{N}(\bar{\beta}_2, \sigma^2)$ .

Writing

$$h(\phi, \beta_1, \beta_2, x) = \frac{\phi}{1 + \exp\left(-\frac{x - \beta_1}{\beta_2}\right)},$$

the complete likelihood of the exponentialized model can then be written as:

$$\begin{aligned} f(y, \phi, \beta_1, \beta_2, \theta) &= 2\pi\sigma^2(2\pi\sigma_\varepsilon^2)^{-nm/2}(2\pi\tau^2)^{-n/2} \cdot \exp\left[-\frac{1}{2\sigma_\varepsilon^2} \sum_{i,j} \left(\frac{y_{i,j}}{h(\phi_i, \beta_1, \beta_2, x_j)} - 1\right)^2\right. \\ &\quad \left. - \sum_{i,j} \log(h(\phi_i, \beta_1, \beta_2, x_j)) - \sum_i \frac{(\phi_i - \mu)^2}{2\tau^2} - \frac{(\beta_1 - \bar{\beta}_1)^2}{2\sigma^2} - \frac{(\beta_2 - \bar{\beta}_2)^2}{2\sigma^2}\right], \end{aligned}$$

where  $\theta = (\mu, \bar{\beta}_1, \bar{\beta}_2)$  are the exponentialized model parameters to estimate.

**Remark 4.1.** It would be easy to suppose  $\tau$  and  $\sigma_\varepsilon^2$  unknown and estimate them using the SAEM algorithm. Those parameters would leave the joint distribution curved exponential and it would not be necessary to further exponentialize the model. To simplify, we assume them known here.

It is then easy to show that this likelihood belongs to the curved exponential family with sufficient statistics being:

$$\begin{cases} S_1(\phi) = \sum_i \phi_i, \\ S_2(\beta_1) = \beta_1, \\ S_3(\beta_2) = \beta_2. \end{cases}$$

The maximum likelihood estimator can then be expressed as a function of  $S_1(\phi)$ ,  $S_2(\beta_1)$  and  $S_3(\beta_2)$  as follows:

$$\begin{cases} \hat{\mu} = S_1(\phi)/n, \\ \hat{\beta}_1 = S_2(\beta_1), \\ \hat{\beta}_2 = S_3(\beta_2). \end{cases}$$

Because we cannot easily sample  $(\phi, \beta_1, \beta_2)$  from the conditional distribution, we will not directly use the SAEM algorithm but the SAEM-MCMC algorithm. We replace the sampling step by one iteration of a Metropolis Hastings algorithm targeting the posterior distribution. Under hypotheses presented in [15], this process converges towards the same limit as the SAEM algorithm. In particular, it has been proved in [16] that those conditions are indeed verified here and thus that the algorithm converges. Moreover, as the limit is the same than the one given by the SAEM, our Theorem 3.5 still applies.

We then create a synthetic dataset of a thousand observations following this model (100 subjects observed at 10 different ages). Knowing the exact value of  $\mu$ , we plot the incomplete likelihood of the non-exponentialized model  $g_{NE}$  as a function of  $(\beta_1, \beta_2)$  (Fig. 4a). We also plot its behaviour around the maximum and along the axes  $\beta_1$  and  $\beta_2$  (Figs. 4b and 4c).

As we can see, the function is not symmetric around the maximum. Hence, there should be a bias while estimating the maximum likelihood using the exponentialized model. More precisely, we can see the error in  $\beta_2$  should be larger than the one in  $\beta_1$  as the function is less symmetric along the  $y$  axis than along the  $x$  axis.

To verify this heuristic, we use the SAEM-MCMC algorithm and launch our algorithm a hundred times for different values of  $\sigma$ . We then compare the results given by the SAEM-MCMC algorithm to the exact value of the maximum likelihood of the initial model. Because we know the exact parameters from which the dataset has been simulated, we are also able to compute numerically the solution of the equation (3.1) as a function of  $\sigma$ . The results are presented in Figure 5.

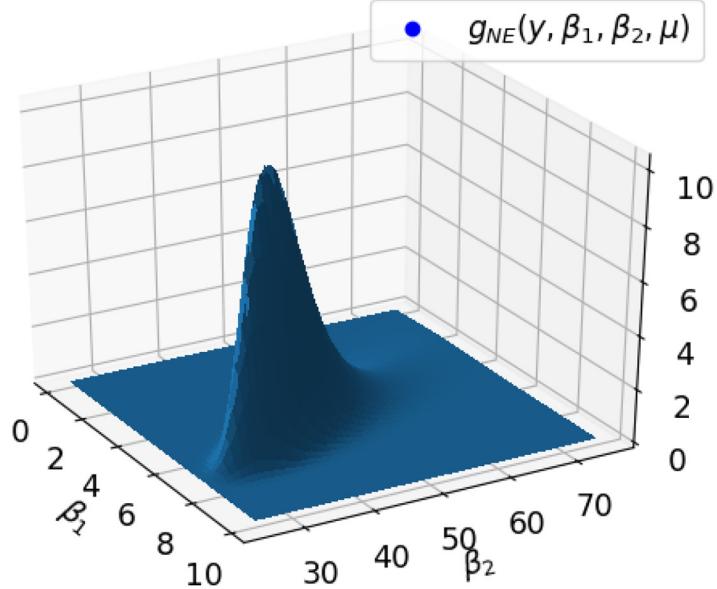
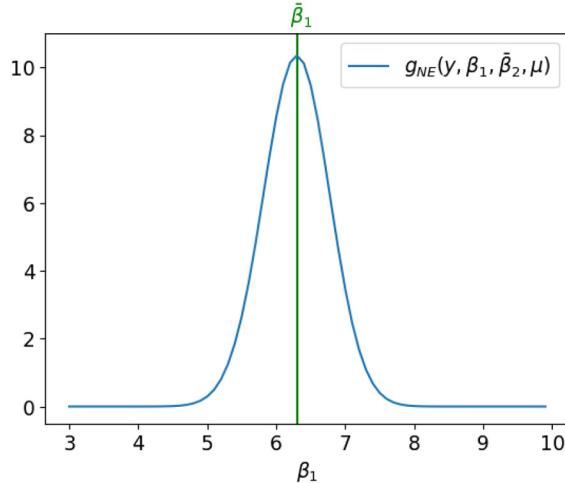
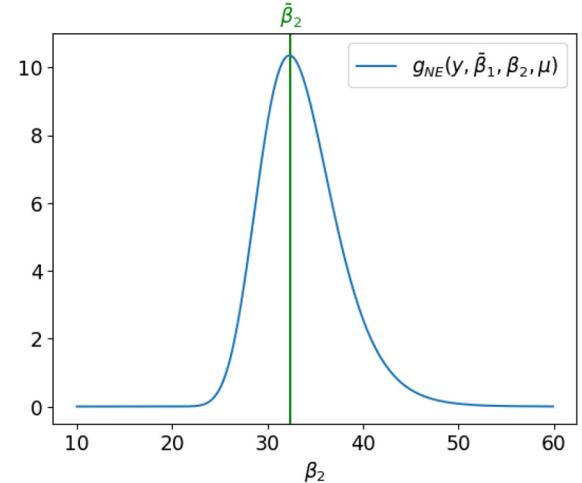
For  $\sigma \geq 1$ , the results of the simulation follow our theory with the estimated parameters estimated close to the solution of the equation (3.1). Moreover, as expected, the error is bigger in the estimation of  $\beta_2$  than in the estimation of  $\beta_1$  (see axis scale).

However, for a small  $\sigma$ , the algorithm does not converge. Indeed, in that case, the variance of the conditional distribution is really small as it is proportional to  $\exp\left(-\frac{(\beta-\bar{\beta})^2}{2\sigma^2}\right)$ . In fact, although the algorithm converges theoretically, numerically, the time of convergence is too long to be achieved, and the returned parameters stay near their initial values ( $\beta_1 = 6$ ,  $\beta_2 = 34$  here).

## 4.2. Proposition of a new algorithm

To prevent this phenomenon, we now propose a transformation of the algorithm that will allow a better estimation of the real maximum of the non-exponentialized likelihood. We will still use the exponential trick but using an adaptive  $\sigma$  along the iterations. The goal is to allow the estimate to escape from its initial value while converging towards a point closer to the true maximum.

We propose to first run the algorithm with  $\sigma = 1$  for a certain number  $k_\sigma$  of iterations and then reduce the value of  $\sigma$  by multiplying it by 0.9. We iterate this process  $p$  times, with  $p$  fixed, every  $k_\sigma$  iterations. We then let the algorithm converge with this small, now fixed, value of  $\sigma = \sigma_{end}$ . This may be seen as launching the algorithm several times with an initialization closer and closer to the true maximum likelihood. In other words,

(a) Plot of the incomplete likelihood of the initial model as a function of  $(\beta_1, \beta_2)$ .(b) Plot of the incomplete likelihood of the initial model as a function of  $\beta_1$  for  $\bar{\beta}_2$  the argmax of likelihood.(c) Plot of the incomplete likelihood of the initial model as a function of  $\beta_2$  for  $\bar{\beta}_1$  the argmax of likelihood.FIGURE 4. Plot of the incomplete likelihood of the initial model as a function of  $(\beta_1, \beta_2)$  along different Sections for  $\mu = 5$ .

it consists in reducing the value of  $\sigma$  along iterations until a certain fixed small value  $\sigma_{end}$  is reached. At this point, we again have a classical MCMC-SAEM but with a clever initialization and a small sigma that allows a quicker convergence. In particular, the results of the previous theorem can still be applied with variance  $\sigma_{end}$ .

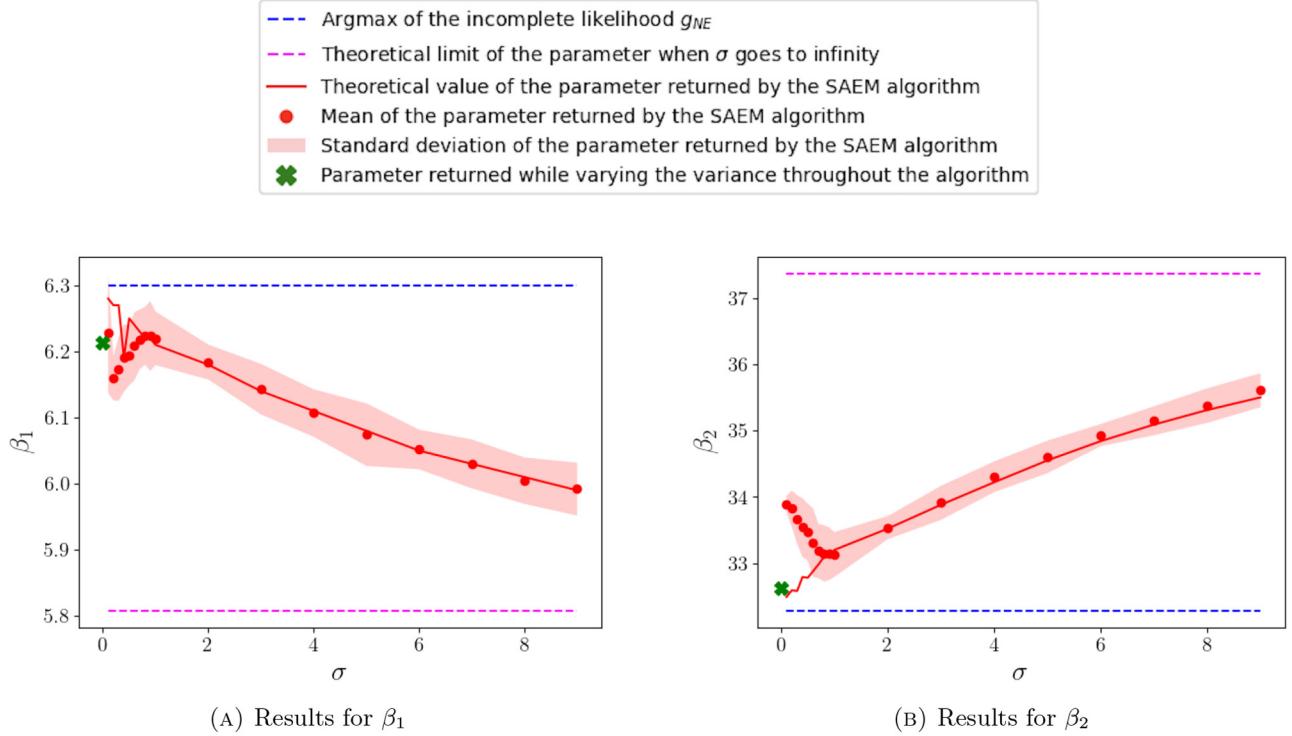


FIGURE 5. The red line represents the theoretical value towards which the algorithm is supposed to converge. The red points are the means of the parameters estimated over 100 iterations with their standard deviations represented by the red zone. In dotted blue is the maximum likelihood of the initial model. In magenta, the theoretical limit towards which the parameter converges when  $\sigma$  goes to infinity. Finally, the green cross represents the value returned while varying the variance of the exponentialized model throughout the algorithm.

While the algorithm will not converge towards the real maximum likelihood estimate as  $\sigma_{end}$  is still positive during the last iterations, the error should be smaller than before as the variance has been significantly reduced.

This transformation can be compared to a simulated annealing, as implemented, for instance in the Monolix software [22]. Here however, we change the model by modifying only the law of the prior  $\psi$ , whereas, with simulated annealing, the whole log conditional distribution is divided by the temperature. The two processes could still be used in conjunction, either with only one temperature parameter or with a temperature parameter and another parameter controlling the decrease of  $\sigma$ .

To test this new algorithm, we launch this process a hundred times. We choose  $k_\sigma = 1000$  and reduce the value of  $\sigma$  from 1 to 0.04. Once this value of  $\sigma = 0.04$  is reached, we let the algorithm converge. We present the means and variances of the estimated parameters in Table 1 and as green crosses in Figure 5. If we do not reach the maximum likelihood of the initial model, the error for  $\beta_2$  is now smaller: 1.04% while it was at least 2.6% without reducing the variance throughout the algorithm. As for  $\beta_1$ , the error is of the same order as before.

**Remark 4.2.** Here, the value of  $k_\sigma$  and the final value of  $\sigma$  are fixed in advance. One could try to automatically choose  $k_\sigma$  along the algorithm by checking the variation of the estimated parameters along the iterations and reduce the value of  $\sigma$  when those parameters begin to oscillate around the limit parameter  $\bar{\psi}_\sigma$ . Similarly, one could choose to reduce the value of  $\sigma$  until the values of  $\bar{\psi}_\sigma$  stay similar from one  $\sigma$  to another.

TABLE 1. Mean and variance of the parameters estimated while reducing the variance throughout the algorithm. To be compared with the maximum likelihood of the non-exponentialized model reached for  $\beta_1 = 6.3$  and  $\beta_2 = 32.28$ .

Mean of $\bar{\beta}_1$	Variance of $\bar{\beta}_1$	Mean of $\bar{\beta}_2$	Variance of $\bar{\beta}_2$
6.21	0.10	32.62	0.26

**Remark 4.3.** In [16], the authors use this model and algorithm on a real dataset for different values of  $\sigma$ . They conclude that the estimation of  $(\beta_1, \beta_2)$  does not seem to depend on the choice of  $\sigma$ . In fact, for the particular values of this real dataset, the likelihood is practically symmetric around its maximum. Hence, the error made in that case is indeed small for any  $\sigma$ .

## 5. CONCLUSION

In this paper, we have proved that the exponentialization process does not converge in general towards the maximum likelihood of the initial model using the SAEM or SAEM-MCMC algorithm. If the error converges towards 0 when  $\sigma$  goes to 0, it is numerically impossible to take  $\sigma$  too small as the algorithm is numerically never able to converge. To overcome this problem, we propose a new numerical scheme consisting in launching the algorithm several times while making the variance of the exponentialized model decrease. Thanks to our theoretical results, we show that this new process converges towards a better estimation of the maximum of likelihood of the initial model, as verified by the numerical simulations. Hence, we are able to approach the exact maximum likelihood even in the case where our likelihood does not belong to the curved exponential family.

## FUNDING

This work was supported in part the French government under management of Agence Nationale de la Recherche as part of the “Investissements d’avenir” program, reference ANR19-P3IA-0001 (PRAIRIE 3IA Institut).

## APPENDIX A. PROOF OF THEOREM 2 FOR $m \geq 2$

*Proof.* We prove the Theorem 2 in the case  $m \geq 2$  and  $l = 2$ . For  $l \geq 3$ , the proof could be obtained using Taylor Lagrange formula at a higher order.

We recall the following equation verified by the limit (Thm. 1):

$$\int_{\mathbb{R}^m} v_k g(\theta_\infty, \bar{\psi}_\sigma + v) \exp\left(-\frac{\|v\|^2}{2\sigma^2}\right) dv = 0, \quad (\text{A.1})$$

### A.1 First step: $d(\bar{\psi}_\sigma, \mathcal{L}) \xrightarrow[\sigma \rightarrow 0]{} 0$

We suppose that  $d(\bar{\psi}_\sigma, \mathcal{L})$  does not converge towards 0. Even if it means extracting a subsequence, we can suppose that,  $\exists c > 0, \forall \sigma > 0, d(\bar{\psi}_\sigma, \mathcal{L}) > 3c$ . As for the one-dimensional proof, we forget the  $\theta$  in  $g$  and write  $g(\psi) = g(\theta_\infty, \psi)$ . We also set  $K_0 = K \setminus \{y \mid d(y, \mathcal{L}) < c\}$ .

We want to show that

$$\exists c_0 > 0, \exists c_1 > 0, \forall y \in K_0, \exists 1 \leq i \leq m, \forall x \text{ verifying } \|x - y\| \leq c_1, \left| \frac{\partial g}{\partial \psi_i}(x) \right| > c_0. \quad (\text{A.2})$$

By contradiction, we can take  $c_0 = 1/n$  and extract a converging subsequence in the compact  $K_0$  to find:

$$\forall c_1 > 0, \exists y \in K_0, \forall 1 \leq i \leq m, \exists x \text{ verifying } \|x - y\| \leq c_1, \left| \frac{\partial g}{\partial \psi_i}(x) \right| = 0.$$

However, because  $y \notin \mathcal{L}$ , there exists  $1 \leq j \leq m$  such that  $\left| \frac{\partial g}{\partial \psi_j}(y) \right| \neq 0$ . If we take  $c_1$  small enough then, for all  $x$  such that  $\|y - x\| \leq c_1$ ,  $\left| \frac{\partial g}{\partial \psi_j}(x) \right| \neq 0$  and we find a contradiction. Hence, the condition (A.2) is verified.

Hence,

$$\exists c_0 > 0, \exists c_1 > 0, \forall \sigma > 0, \exists k \in [|1, m|], \forall v \text{ verifying } \|v - \bar{\psi}_\sigma\| \leq c_1, \left| \frac{\partial g}{\partial \psi_k}(v) \right| > c_0.$$

As for the proof in dimension 1, we split up our integral in two parts:  $I_1 = \{v \mid \forall i \in [|1, n|], v_i \leq c_2\}$  and  $I_2 = \{v \mid \exists i \in [|1, n|], v_i \geq c_2\}$  where  $c_2$  is chosen such that  $\{v \mid \forall i \in [|1, n|], v_i \leq c_2\} \subset \{v \mid \|v\| \leq c_1\}$ .

First, on  $I_2$ ,

$$\begin{aligned} \int_{I_2} v_k g(\bar{\psi}_\sigma + v) \exp\left(-\frac{\|v\|^2}{2\sigma^2}\right) dv &\geq \int_{I_2, v_k \leq 0} v_k g(\bar{\psi}_\sigma + v) \exp\left(-\frac{\|v\|^2}{2\sigma^2}\right) dv \\ &\geq \|g\|_\infty \left( \int_{\mathbb{R}^m, v_k \leq 0} v_k \exp\left(-\frac{\|v\|^2}{2\sigma^2}\right) dv - \int_{I_1, v_k \leq 0} v_k \exp\left(-\frac{\|v\|^2}{2\sigma^2}\right) dv \right) \\ &\geq -\sigma^2 (\sqrt{2\pi}\sigma)^{m-1} \|g\|_\infty \left( 1 - \left( 1 - \exp\left(-\frac{c_2^2}{2\sigma^2}\right) \right) \operatorname{erf}\left(\frac{c_2}{\sqrt{2}\sigma}\right)^{m-1} \right) \end{aligned}$$

where  $\operatorname{erf}$  is the error function defined by  $\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$ .

We now integrate on  $I_1$ . We write  $v^{-k}$  the vector such that,  $\forall i \neq k, (v^{-k})_i = v_i$  and  $(v^{-k})_k = -v_k$ . Then, using the mean value theorem, we have

$$\begin{aligned} \int_{I_1} v_k g(\bar{\psi}_\sigma + v) \exp\left(-\frac{\|v\|^2}{2\sigma^2}\right) dv &= \int_{I_1, v_k \leq 0} v_k (g(\bar{\psi}_\sigma + v) - g(\bar{\psi}_\sigma + v^{-k})) \exp\left(-\frac{\|v\|^2}{2\sigma^2}\right) dv \\ &= \int_{I_1} 2v_k^2 \frac{\partial g}{\partial \psi_k}(\tilde{\psi}_v) \exp\left(-\frac{\|v\|^2}{2\sigma^2}\right) dv \end{aligned}$$

where, for  $i \neq k$ ,  $(\tilde{\psi}_v)_i = (\bar{\psi}_\sigma)_i + v_i$  and  $(\tilde{\psi}_v)_k \in [(\bar{\psi}_\sigma)_k - v_k, (\bar{\psi}_\sigma)_k + v_k]$ . But we know that  $\frac{\partial g}{\partial \psi_k}$  does not cancel on  $I_1$ . Hence, it is either positive or negative. If it is positive, we find:

$$\begin{aligned} \int_{I_1} v_k g(\bar{\psi}_\sigma + v) \exp\left(-\frac{\|v\|^2}{2\sigma^2}\right) dv &\geq 2c_0 \int_{I_1, v_k \leq 0} v_k^2 \exp\left(-\frac{\|v\|^2}{2\sigma^2}\right) dv \\ &\geq 2c_0 (\sqrt{2\pi}\sigma)^{m-1} \sigma^2 \left[ -c_2 \exp\left(-\frac{c_2^2}{2\sigma^2}\right) + \sigma \frac{\sqrt{\pi}}{2} \operatorname{erf}\left(\frac{c_2}{\sqrt{2}\sigma}\right) \right] \operatorname{erf}\left(\frac{c_2}{\sqrt{2}\sigma}\right)^{m-1} \end{aligned}$$

Finally, using (A.1), we have:

$$\begin{aligned} 2c_0 \left[ -\frac{c_2}{\sigma} \exp\left(-\frac{c_2^2}{2\sigma^2}\right) + \frac{\sqrt{\pi}}{2} \operatorname{erf}\left(\frac{c_2}{\sqrt{2}\sigma}\right) \right] \operatorname{erf}\left(\frac{c_2}{\sqrt{2}\sigma}\right)^{m-1} \\ \leq \|g\|_\infty \left[ \frac{1 - \operatorname{erf}\left(\frac{c_2}{\sqrt{2}\sigma}\right)^{m-1}}{\sigma} + \frac{1}{\sigma} \exp\left(-\frac{c_2^2}{2\sigma^2}\right) \operatorname{erf}\left(\frac{c_2}{\sqrt{2}\sigma}\right)^{m-1} \right] \end{aligned} \quad (\text{A.3})$$

But,

$$\operatorname{erf}(x) =_{x \rightarrow \infty} 1 - \frac{\exp(-x^2)}{\sqrt{\pi}x} + o\left(\frac{\exp(-x^2)}{x}\right).$$

Hence, when  $\sigma$  goes to 0, the left-hand side of the inequality goes to  $\sqrt{\pi}c_0$  while the right-hand side goes to 0. We thus find a contradiction.

Finally, if  $\frac{\partial g}{\partial \psi_k}$  is not positive on  $I_1$  (as supposed here) but negative, we can use the same method to find an upper bound on the integral on  $I_2$  and on the integral on  $I_1$ . We would then find the same inequality as in (A.3).

Hence, in all cases, we have proved that  $d(\bar{\psi}_\sigma, \mathcal{L}) \xrightarrow[\sigma \rightarrow 0]{} 0$ .

## A.2 Second step: Choice of the basis

The upper bound using second derivatives is more complex to obtain for  $m > 1$  as crossed partial derivatives appear that can be either positive or negative. To control those parts, the choice of the compact is more complex. We will first show that we can express our vector  $v$  and our function  $g$  in any orthonormal basis and still have the equation (A.1).

Indeed, let  $P$  be a change-of-basis matrix. Then, because the equation (A.1) is linear on  $v_k$  and true for all  $k \in [|1, m|]$ , we still have:

$$\int_{\mathbb{R}^m} (Pv)_k g(\bar{\psi}_\sigma + v) \exp\left(-\frac{\|v\|^2}{2\sigma^2}\right) dv = 0.$$

Using the change of variable  $u = Pv$ , we then find, for any  $k \in [|1, m|]$ :

$$\int_{\mathbb{R}^m} u_k g(P^{-1}(P\bar{\psi}_\sigma + u)) \exp\left(-\frac{\|u\|^2}{2\sigma^2}\right) dv = 0.$$

We write  $h : u \mapsto g(P^{-1}u)$ . Hence,  $h$  verifies the equation (A.1).

We can thus choose to express our function  $g$  in any base. In particular, we write  $\psi_M$  the nearest maximum of  $\bar{\psi}_\sigma$ . Then, the Hessian of  $g$  at  $\psi_M$  is a negative symmetric matrix. Hence, it is diagonal in an orthonormal basis. We choose to express  $g$  in that basis. With a change of notation, we can hence assume that the hessian of  $g$  at  $\psi_M$  is diagonal. In particular, for all  $i \neq j \in [|1, m|]$ ,

$$\frac{\partial^2 g}{\partial \psi_i^2}(\psi_M) < 0 \text{ and } \frac{\partial^2 g}{\partial \psi_i \partial \psi_j}(\psi_M) = 0.$$

In particular, we are now able to impose a condition between the second derivatives of  $g$  on a compact centered around  $\psi_M$ . There exists  $K_0$  compact such that,

$$\begin{aligned} -\sup_{K_0} \frac{\partial^2 g}{\partial \psi_k^2} &> (m-1) \sup \left\{ \frac{\partial^2 g}{\partial \psi_k \partial \psi_j}(v) \mid v \in K_0, j \neq k, \frac{\partial^2 g}{\partial \psi_k \partial \psi_j}(v) > 0 \right\} \\ &\quad - \frac{m-1}{2} \inf \left\{ \frac{\partial^2 g}{\partial \psi_k \partial \psi_j}(v) \mid v \in K_0, j \neq k, \frac{\partial^2 g}{\partial \psi_k \partial \psi_j}(v) < 0 \right\} \end{aligned} \quad (\text{A.4})$$

### A.3 Third step: Search of the upper bound

As for the proof in 1D, we will split our integral into two parts and say that neither can be too big for the complete integral to be equal to 0. More precisely, for  $\sigma > 0$ , let  $k$  be the coordinate such that  $|(\bar{\psi}_\sigma)_k - (\psi_M)_k| = \max |(\bar{\psi}_\sigma)_i - (\psi_M)_i|$ . We write for  $i \in [|1, m|]$ ,  $(\alpha_\sigma)_i = |(\bar{\psi}_\sigma)_i - (\psi_M)_i|$ . The goal is to show that  $(\alpha_\sigma)_k$  goes to 0 when  $\sigma$  goes to 0.

Let  $c > 0$  and  $\sigma$  small enough such that

$$I_1 := \{v \in \mathbb{R}^m \mid v_k \in [-(\alpha_\sigma)_k, (\alpha_\sigma)_k] \text{ and, for } i \neq k, v_i \in [-c, c]\} \subset K_0.$$

On  $I_1^c$ , we use the same upper bounds as in the first step to find:

$$\begin{aligned} \int_{I_1^c} v_k g(\bar{\psi}_\sigma + v) \exp\left(-\frac{\|v\|^2}{2\sigma^2}\right) dv &\geq -\sigma^2 (\sqrt{2\pi}\sigma)^{m-1} \|g\|_\infty \\ &\cdot \left( 1 - \left( 1 - \exp\left(-\frac{(\alpha_\sigma)_k^2}{2\sigma^2}\right) \right) \operatorname{erf}\left(\frac{c}{\sqrt{2}\sigma}\right)^{m-1} \right) \end{aligned} \quad (\text{A.5})$$

On  $I_1$  we will use once again the mean value theorem, first between  $\bar{\psi}_\sigma + v^{-k}$  and  $\bar{\psi}_\sigma + v$  to find  $\tilde{\psi}_v \in K_0$  such that, for  $i \neq k$ ,  $(\tilde{\psi}_v)_i = (\bar{\psi}_\sigma + v)_i$ ,  $(\tilde{\psi}_v)_k \in [(\bar{\psi}_\sigma - v)_k, (\bar{\psi}_\sigma + v)_k]$  and

$$g(\bar{\psi}_\sigma + v) - g(\bar{\psi}_\sigma + v^{-k}) = 2v_k \frac{\partial g}{\partial \psi_k}(\tilde{\psi}_v)$$

and then between  $\tilde{\psi}_v$  and  $\psi_M$  to find  $\tilde{\psi}_v^1 \in K_0$  such that:

$$\frac{\partial g}{\partial \psi_k}(\tilde{\psi}_v) = \sum_{i=1}^m (\tilde{\psi}_v - \psi_M)_i \frac{\partial^2 g}{\partial \psi_k \partial \psi_i}(\tilde{\psi}_v^1).$$

Even if it means changing basis, we can assume that,  $\forall i \in [|1, m|]$ ,  $(\alpha_\sigma)_i = |(\bar{\psi}_\sigma)_i - (\psi_M)_i| = (\psi_M)_i - (\bar{\psi}_\sigma)_i$  without modifying the hypothesis (A.4).

The difficulty to find upper bounds is that  $\frac{\partial^2 g}{\partial \psi_k \partial \psi_i}(\tilde{\psi}_v^1)$  and also  $(\tilde{\psi}_v - \psi_M)_i$  can be either positive or negative.

Using those previous equalities and the facts that  $\frac{\partial^2 g}{\partial \psi_k^2} < 0$  on  $K_0$  and  $(\tilde{\psi}_v - \psi_M)_k \leq (\bar{\psi}_\sigma + v - \psi_M)_k = (v - \alpha_\sigma)_k$ , we have:

$$\int_{I_1} v_k g(\bar{\psi}_\sigma + v) \exp\left(-\frac{\|v\|^2}{2\sigma^2}\right) dv \geq 2 \int_{I_1, v_k \geq 0} v_k^2 \sum_{i=1}^m (v - \alpha_\sigma)_i \frac{\partial^2 g}{\partial \psi_k \partial \psi_i}(\tilde{\psi}_v^1) \exp\left(-\frac{\|v\|^2}{2\sigma^2}\right) dv$$

We will study the different terms of the sum differently according to  $i = k$ , or  $i \neq k$ .

First, for  $i = k$ , using the fact that  $(v - \alpha_\sigma)_k \leq 0$  on  $I_1$ , we can compute the integral using integration per part and the function erf defined above to find:

$$\begin{aligned} \int_{I_1, v_k \geq 0} 2v_k^2(v - \alpha_\sigma)_k \frac{\partial^2 g}{\partial \psi_k^2}(\tilde{\psi}_v^1) \exp\left(-\frac{\|v\|^2}{2\sigma^2}\right) dv &\geq \sup_{K_0} \left( \frac{\partial^2 g}{\partial \psi_k^2} \right) \int_{I_1, v_k \geq 0} 2v_k^2(v - \alpha_\sigma)_k \exp\left(-\frac{\|v\|^2}{2\sigma^2}\right) dv \\ &= -2 \sup_{K_0} \left( \frac{\partial^2 g}{\partial \psi_k^2} \right) (\sqrt{2\pi}\sigma)^{m-1} \operatorname{erf}\left(\frac{c}{\sqrt{2}\sigma}\right)^{m-1} (\alpha_\sigma)_k^4 \left[ \sqrt{\frac{\pi}{2}} \left( \frac{\sigma}{(\alpha_\sigma)_k} \right)^3 \operatorname{erf}\left(\frac{(\alpha_\sigma)_k}{\sqrt{2}\sigma}\right) \right. \\ &\quad \left. - 2 \left( \frac{d}{(\alpha_\sigma)_k} \right)^4 \left( 1 - \exp\left(-\frac{(\alpha_\sigma)_k}{2\sigma^2}\right) \right) \right] \end{aligned}$$

For  $i \neq k$ , we do similar computations remarking that:

– if  $\frac{\partial^2 g}{\partial \psi_k \partial \psi_i}(\tilde{\psi}_v^1) > 0$  and  $(v - \alpha_\sigma)_i > 0$ ,

$$2v_k^2(v - \alpha_\sigma)_i \frac{\partial^2 g}{\partial \psi_k \partial \psi_i}(\tilde{\psi}_v^1) \exp\left(-\frac{\|v\|^2}{2\sigma^2}\right) > 0$$

– if  $\frac{\partial^2 g}{\partial \psi_k \partial \psi_i}(\tilde{\psi}_v^1) > 0$  and  $(v - \alpha_\sigma)_i < 0$ , with  $K_0^+ = \left\{ v \in K_0 \mid \frac{\partial^2 g}{\partial \psi_k \partial \psi_i}(v) > 0 \right\}$ ,

$$2v_k^2(v - \alpha_\sigma)_i \frac{\partial^2 g}{\partial \psi_k \partial \psi_i}(\tilde{\psi}_v^1) \exp\left(-\frac{\|v\|^2}{2\sigma^2}\right) > 2v_k^2(v - \alpha_\sigma)_i \exp\left(-\frac{\|v\|^2}{2\sigma^2}\right) \sup_{K_0^+, i} \frac{\partial^2 g}{\partial \psi_k \partial \psi_i}$$

– if  $\frac{\partial^2 g}{\partial \psi_k \partial \psi_i}(\tilde{\psi}_v^1) < 0$  and  $(v - \alpha_\sigma)_i < 0$

$$2v_k^2(v - \alpha_\sigma)_i \frac{\partial^2 g}{\partial \psi_k \partial \psi_i}(\tilde{\psi}_v^1) \exp\left(-\frac{\|v\|^2}{2\sigma^2}\right) > 0$$

– if  $\frac{\partial^2 g}{\partial \psi_k \partial \psi_i}(\tilde{\psi}_v^1) < 0$  and  $(v - \alpha_\sigma)_i > 0$ , with  $K_0^- = \left\{ v \in K_0 \mid \frac{\partial^2 g}{\partial \psi_k \partial \psi_i}(v) < 0 \right\}$ ,

$$2v_k^2(v - \alpha_\sigma)_i \frac{\partial^2 g}{\partial \psi_k \partial \psi_i}(\tilde{\psi}_v^1) \exp\left(-\frac{\|v\|^2}{2\sigma^2}\right) > 2v_k^2(v - \alpha_\sigma)_i \exp\left(-\frac{\|v\|^2}{2\sigma^2}\right) \inf_{K_0^-, i} \frac{\partial^2 g}{\partial \psi_k \partial \psi_i}$$

Hence, for  $i \neq k$ , we write  $I_1^- = \{v \in I_1 \mid v_k \geq 0, v_i \leq (\alpha_\sigma)_i\}$  and:

$$\begin{aligned} \int_{I_1^-} 2v_k^2(v - \alpha_\sigma)_i \frac{\partial^2 g}{\partial \psi_k \partial \psi_i}(\tilde{\psi}_v^1) \exp\left(-\frac{\|v\|^2}{2\sigma^2}\right) dv &\geq \sup_{K_0^+, i} \frac{\partial^2 g}{\partial \psi_k \partial \psi_i} \int_{I_1^-} 2v_k^2(v - \alpha_\sigma)_i \exp\left(-\frac{\|v\|^2}{2\sigma^2}\right) dv \\ &= 2 \sup_{K_0^+, i} \frac{\partial^2 g}{\partial \psi_k \partial \psi_i} (\sqrt{2\pi}\sigma)^{m-1} \operatorname{erf}\left(\frac{c}{\sqrt{2}\sigma}\right)^{m-1} (\alpha_\sigma)_k^4 \\ &\quad \cdot \left[ \sqrt{\frac{\pi}{2}} \left( \frac{\sigma}{(\alpha_\sigma)_k} \right)^4 \operatorname{erf}\left(\frac{(\alpha_\sigma)_k}{\sqrt{2}\sigma}\right) - \left( \frac{\sigma}{(\alpha_\sigma)_k} \right)^3 \exp\left(-\frac{(\alpha_\sigma)_k^2}{2\sigma^2}\right) \right] \\ &\quad \cdot \left[ \frac{\exp\left(\frac{-c^2}{2\sigma^2}\right) - \exp\left(\frac{-(\alpha_\sigma)_j^2}{2\sigma^2}\right)}{\sqrt{2\pi} \operatorname{erf}\left(\frac{c}{\sqrt{2}\sigma}\right)} - \frac{1}{2} \frac{(\alpha_\sigma)_j}{d} \left( 1 + \frac{\operatorname{erf}\left(\frac{(\alpha_\sigma)_j}{\sqrt{2}\sigma}\right)}{\operatorname{erf}\left(\frac{c}{\sqrt{2}\sigma}\right)} \right) \right] \end{aligned}$$

Similarly, with  $I_1^+ = \{v \in I_1 \mid v_k \geq 0, v_i \geq (\alpha_\sigma)_i\}$ ,

$$\begin{aligned} \int_{I_1^+} 2v_k^2(v - \alpha_\sigma)_i \frac{\partial^2 g}{\partial \psi_k \partial \psi_i}(\tilde{\psi}_v^1) \exp\left(-\frac{\|v\|^2}{2\sigma^2}\right) dv &\geq \inf_{K_0^-, i} \frac{\partial^2 g}{\partial \psi_k \partial \psi_i} \int_{I_1^+} 2v_k^2(v - \alpha_\sigma)_i \exp\left(-\frac{\|v\|^2}{2\sigma^2}\right) dv \\ &= 2 \inf_{K_0^-, i} \frac{\partial^2 g}{\partial \psi_k \partial \psi_i}(\sqrt{2\pi}\sigma)^{m-1} \operatorname{erf}\left(\frac{c}{\sqrt{2}\sigma}\right)^{m-1} (\alpha_\sigma)_k^4 \\ &\quad \cdot \left[ \sqrt{\frac{\pi}{2}} \left( \frac{\sigma}{(\alpha_\sigma)_k} \right)^4 \operatorname{erf}\left(\frac{(\alpha_\sigma)_k}{\sqrt{2}\sigma}\right) - \left( \frac{\sigma}{(\alpha_\sigma)_k} \right)^3 \exp\left(-\frac{(\alpha_\sigma)_k^2}{2\sigma^2}\right) \right] \\ &\quad \cdot \left[ \frac{\exp\left(\frac{-(\alpha_\sigma)_j^2}{2\sigma^2}\right) - \exp\left(\frac{-c^2}{2\sigma^2}\right)}{\sqrt{2\pi} \operatorname{erf}\left(\frac{c}{\sqrt{2}\sigma}\right)} - \frac{1}{2} \frac{(\alpha_\sigma)_j}{d} \left( 1 - \frac{\operatorname{erf}\left(\frac{(\alpha_\sigma)_j}{\sqrt{2}\sigma}\right)}{\operatorname{erf}\left(\frac{c}{\sqrt{2}\sigma}\right)} \right) \right] \end{aligned}$$

Those three upper bounds are quite complex, but we can remark that they can be written as  $d^{m-1} \operatorname{erf}\left(\frac{c}{\sqrt{2}\sigma}\right)^{m-1} (\alpha_\sigma)_k^4 h\left(\frac{(\alpha_\sigma)_k}{\sigma}\right)$ .

We will now see that this function  $h$  is strictly positive at infinity.

Indeed, using all the previous upper bounds presented previously, equation (A.1) and using the fact that  $\operatorname{erf}\left(\frac{c}{\sqrt{2}\sigma}\right) \geq 1/2$  for  $\sigma$  small enough, we can write:

$$(\alpha_\sigma)_k^4 \operatorname{erf}\left(\frac{c}{\sqrt{2}\sigma}\right)^{m-1} \left( h\left(\frac{\alpha_k}{\sigma}\right) - \frac{1 - \operatorname{erf}(c/\sqrt{2}\sigma)}{\sigma^2 \operatorname{erf}(c/\sqrt{2}\sigma)} \right) \leq \frac{\|g\|_\infty}{2} \sigma^2$$

with:

$$\begin{aligned} h(x) &= \frac{e^{x^2/2}}{x^4} \left[ -\sup_{K_0} \frac{\partial^2 g}{\partial \psi_k^2} \left( \frac{\sqrt{\pi}}{2} x \operatorname{erf}(x/\sqrt{2}) - 2(1 - e^{-x^2/2}) \right) \right. \\ &\quad + \left( \frac{\sqrt{\pi}}{2} x \operatorname{erf}(x/\sqrt{2}) - x^2 e^{-x^2/2} \right) \cdot \left[ (m-1) \left( \inf_{K_0^-, i} \frac{\partial^2 g}{\partial \psi_k \partial \psi_i} - \sup_{K_0^+, i} \frac{\partial^2 g}{\partial \psi_k \partial \psi_i} \right) \left( \frac{2}{\sqrt{2\pi}x} + \frac{1}{2} \right) \right. \\ &\quad \left. \left. - \frac{m-1}{2} \sup_{K_0^+, i} \frac{\partial^2 g}{\partial \psi_k \partial \psi_i} \right] \right] \end{aligned}$$

is a function independent of  $\alpha_\sigma$  and  $\sigma$ .

In particular, when  $x$  goes to infinity,  $h(x)$  is equivalent to:

$$\frac{e^{x^2/2}}{x^4} \frac{\sqrt{\pi}}{2} x \left[ -\sup_{K_0} \frac{\partial^2 g}{\partial \psi_k^2} + \frac{m-1}{2} \inf_{K_0^-, i} \frac{\partial^2 g}{\partial \psi_k \partial \psi_i} - (m-1) \sup_{K_0^+, i} \frac{\partial^2 g}{\partial \psi_k \partial \psi_i} \right]$$

But, according to the hypothesis done on the compact  $K_0$ , this is strictly positive.

Hence, there exist  $c_0 > 0, c_1 > 0$  such that, if  $x \geq c_1$ ,  $h(x) > c_0 > 0$ .

We will now suppose that  $(\alpha_\sigma)_k \geq c_1 \sigma$ . Then,  $h\left(\frac{(\alpha_\sigma)_k}{\sigma}\right) \geq c_0 > 0$ . Moreover,

$$\frac{1 - \operatorname{erf}(c/\sqrt{2}\sigma)}{\sigma^2 \operatorname{erf}(c/\sqrt{2}\sigma)} \xrightarrow[\sigma \rightarrow 0]{} 0.$$

So, for  $\sigma$  small enough, it is smaller than  $c_0/2$  and we finally find:

$$(\alpha_\sigma)_k^4 \leq c_0 \|g\|_\infty \frac{\sigma^2}{\operatorname{erf}\left(\frac{c}{\sqrt{2}\sigma}\right)^{m-1}}$$

Using the fact that  $\operatorname{erf}\left(\frac{c}{\sqrt{2}\sigma}\right)^{m-1} \xrightarrow[\sigma \rightarrow 0]{} 1$  gives us finally the existence of a constant  $c > 0$  such that

$$d(\bar{\psi}_\sigma, \mathcal{L}) \leq (c_1 \sigma) \vee (c\sqrt{\sigma})$$

which allows us to conclude for  $\sigma$  small enough.

#### A.4 Fourth step: Approximation when $\sigma$ goes to infinity

The last step follows the exact same steps as for  $m = 1$ . It is copied here.

We use again the equation (A.1). For all  $\sigma \in \mathbb{R}$ ,  $\forall 1 \leq k \leq m$ ,

$$\int_{\mathbb{R}^m} v_k g(\bar{\psi}_\sigma + v) \exp\left(-\frac{\|v\|^2}{2\sigma^2}\right) dv = 0$$

Using the change of variable  $\bar{\psi}_\sigma + v$ , we find:

$$(\bar{\psi}_\sigma)_k = \frac{\int_{\mathbb{R}^m} v_k g(v) \exp\left(-\frac{\|v - \bar{\psi}_\sigma\|^2}{2\sigma^2}\right) dv}{\int_{\mathbb{R}^m} g(v) \exp\left(-\frac{\|v - \bar{\psi}_\sigma\|^2}{2\sigma^2}\right) dv}$$

But  $\bar{\psi}_\sigma$  is supposed to stay in a compact so,  $\forall v \in \mathbb{R}^m$ ,  $\exp\left(-\frac{\|v - \bar{\psi}_\sigma\|^2}{2\sigma^2}\right) \xrightarrow[\sigma \rightarrow \infty]{} 1$ . Using the integrability of  $g$  and  $v \mapsto v_k g(v)$ , it is easy to conclude using the dominated convergence theorem.  $\square$

## REFERENCES

- [1] O. Ajmal, L. Duchateau and E. Kuhn, Convergent stochastic algorithm for parameter estimation in frailty models using integrated partial likelihood. Preprint [arXiv:1909.07056](https://arxiv.org/abs/1909.07056) (2019).
- [2] S. Allassonnière and J. Chevallier, A new class of em algorithms. escaping local minima and handling intractable sampling. Preprint (2019).
- [3] S. Allassonnière, E. Kuhn, A. Trouvé, *et al.*, Construction of Bayesian deformable models via a stochastic approximation algorithm: a convergence study. *Bernoulli* **16** (2010) 641–678.
- [4] S. Allassonnière, L. Younes, *et al.*, A stochastic algorithm for probabilistic independent component analysis. *Ann. Appl. Stat.* **6** (2012) 125–160.
- [5] C. Andrieu, É. Moulines and P. Priouret, Stability of stochastic approximation under verifiable conditions. *SIAM J. Control Optim.* **44** (2005) 283–312.
- [6] S. Balakrishnan, M.J. Wainwright, B. Yu, *et al.*, Statistical guarantees for the em algorithm: from population to sample-based analysis. *Ann. Stat.* **45** (2017) 77–120.
- [7] S. Benzekry, C. Lamont, A. Beheshti, A. Tracz, J.M. Ebos, L. Hlatky and P. Hahnfeldt, Classical mathematical models for description and prediction of experimental tumor growth. *PLoS Comput Biol.* **10** (2014) e1003800.
- [8] A. Bône, O. Colliot and S. Durrleman, Learning distributions of shape trajectories from longitudinal datasets: a hierarchical model on a manifold of diffeomorphisms, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2018) 9271–9280.
- [9] S. Chrétien and A.O. Hero, On em algorithms and their proximal generalizations. *ESAIM: Probab. Stat.* **12** (2008) 308–326.
- [10] V. Debavelaere, S. Durrleman and S. Allassonnière, Learning the clustering of longitudinal shape data sets into a mixture of independent or branching trajectories. *Int. J. Comput. Vision* (2020) 1–16.
- [11] B. Delyon, M. Lavielle and E. Moulines, Convergence of a stochastic approximation version of the em algorithm. *Ann. Stat.* (1999) 94–128.

- [12] A.P. Dempster, N.M. Laird and D.B. Rubin, Maximum likelihood from incomplete data via the em algorithm. *J. Roy. Stat. Soc. Ser. B (Methodological)* **39** (1977) 1–22.
- [13] A. Dubois, M. Lavielle, S. Gsteiger, E. Pigeolet and F. Mentré, Model-based analyses of bioequivalence crossover trials using the stochastic approximation expectation maximisation algorithm. *Stat. Med.* **30** (2011) 2582–2600.
- [14] J. Guedj and A.S. Perelson, Second-phase hepatitis c virus RNA decline during Telaprevir-based therapy increases with drug effectiveness: implications for treatment duration. *Hepatology* **53** (2011) 1801–1808.
- [15] E. Kuhn and M. Lavielle, Coupling a stochastic approximation version of em with an MCMC procedure. *ESAIM: Probab. Stat.* **8** (2004) 115–131.
- [16] E. Kuhn and M. Lavielle, Maximum likelihood estimation in nonlinear mixed effects models. *Comput. Stat. Data Anal.* **49** (2005) 1020–1038.
- [17] E. Kuhn, C. Matias and T. Rebafka, Properties of the stochastic approximation em algorithm with mini-batch sampling. *Stat. Comput.* **30** (2020) 1725–1739.
- [18] T. Lartigue, S. Durrleman and S. Allassonnière, Deterministic approximate em algorithm; application to the Riemann approximation em and the tempered em. Preprint [arXiv:2003.10126](https://arxiv.org/abs/2003.10126) (2020).
- [19] M. Lavielle, Mixed effects models for the population approach: models, tasks, methods and tools. CRC Press (2014).
- [20] M. Lavielle and F. Mentré, Estimation of population pharmacokinetic parameters of saquinavir in HIV patients with the Monolix software. *J. Pharmacokinet. Pharmacodyn.* **34** (2007) 229–249.
- [21] F. Lindsten, An efficient stochastic approximation em algorithm using conditional particle filters, in 2013 IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE (2013) 6274–6278.
- [22] Lixoft SAS, Monolix (2020).
- [23] J. Ma, L. Xu and M.I. Jordan, Asymptotic convergence rate of the algorithm for Gaussian mixtures. *Neural Comput.* **12** (2000) 2881–2907.
- [24] X.-L. Meng *et al.*, On the rate of convergence of the ECM algorithm. *Ann. Stat.* **22** (1994) 326–339.
- [25] C. Meza, F. Osorio and R. De la Cruz, Estimation in nonlinear mixed-effects models using heavy-tailed distributions. *Stat. Comput.* **22** (2012) 121–139.
- [26] X. Panhard and A. Samson, Extension of the SAEM algorithm for nonlinear mixed models with 2 levels of random effects. *Biostatistics* **10** (2009) 121–135.
- [27] R.A. Redner and H.F. Walker, Mixture densities, maximum likelihood and the algorithm. *SIAM Rev.* **26** (1984) 195–239.
- [28] A. Samson, M. Lavielle and F. Mentré, Extension of the SAEM algorithm to left-censored data in nonlinear mixed-effects model: application to HIV dynamics model. *Comput. Stat. Data Anal.* **51** (2006) 1562–1574.
- [29] J.-B. Schiratti, S. Allassonnière, O. Colliot and S. Durrleman, Learning spatiotemporal trajectories from manifold-valued longitudinal data. *Adv. Neural Inf. Process. Syst.* (2015) 2404–2412.
- [30] D. Sissoko, C. Laouenan, E. Folkesson, A.-B. M'lebing, A.-H. Beavogui, S. Baize, A.-M. Camara, P. Maes, S. Shepherd, C. Danel, *et al.*, Experimental treatment with Favipiravir for Ebola virus disease (the JIKI trial): a historically controlled, single-arm proof-of-concept trial in guinea. *PLoS Med.* **13** (2016) e1001967.
- [31] P. Tseng, An analysis of the algorithm and entropy-like proximal point methods. *Math. Oper. Res.* **29** (2004) 27–44.
- [32] J. Wang, Em algorithms for nonlinear mixed effects models. *Comput. Stat. Data Anal.* **51** (2007) 3244–3256.
- [33] G.C. Wei and M.A. Tanner, A Monte Carlo implementation of the algorithm and the poor man's data augmentation algorithms. *J. Am. Stat. Assoc.* **85** (1990) 699–704.
- [34] C.J. Wu, On the convergence properties of the algorithm. *Ann. Stat.* (1983) 95–103.

## Subscribe to Open (S2O)

### A fair and sustainable open access model



This journal is currently published in open access under a Subscribe-to-Open model (S2O). S2O is a transformative model that aims to move subscription journals to open access. Open access is the free, immediate, online availability of research articles combined with the rights to use these articles fully in the digital environment. We are thankful to our subscribers and sponsors for making it possible to publish this journal in open access, free of charge for authors.

**Please help to maintain this journal in open access!**

Check that your library subscribes to the journal, or make a personal donation to the S2O programme, by contacting [subscribers@edpsciences.org](mailto:subscribers@edpsciences.org)

More information, including a list of sponsors and a financial transparency report, available at: <https://www.edpsciences.org/en/math-s2o-programme>