



**HAL**  
open science

# On the curved exponential family in the Stochastic Approximation Expectation Maximization Algorithm

Vianney Debavelaere, Stéphanie Allasonnière

► **To cite this version:**

Vianney Debavelaere, Stéphanie Allasonnière. On the curved exponential family in the Stochastic Approximation Expectation Maximization Algorithm. 2021. hal-03128554

**HAL Id: hal-03128554**

**<https://hal.science/hal-03128554>**

Preprint submitted on 2 Feb 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# On the curved exponential family in the Stochastic Approximation Expectation Maximization Algorithm

Vianney Debavelaere · Stéphanie Allasonnière

Received: date / Accepted: date

**Abstract** The Expectation-Maximization (EM) Algorithm 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 classical EM such as the Stochastic Approximation EM (SAEM). This algorithm, however, has the disadvantage to require the joint likelihood to belong to the curved exponential family. This hypothesis is a bottleneck in a lot of practical situations where it is not verified. To overcome this problem, Kuhn and Lavielle (2005) introduce a rewriting of the model which “exponentializes” it. It consists in 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 and stochastic EM algorithms apply. Although often used, there is no guarantee that the estimated mean will be close to the maximum likelihood estimate of the initial model. In this paper, we will quantify the error done in this estimation while considering the exponentialized model instead of the initial one. More precisely, we will show that this error tends to 0 as the variance of the new Gaussian distribution goes to 0 while computing an upper bound. By verifying those results on an example, we will see that a compromise must be made in the choice of the variance between the speed of convergence and the tolerated error. Finally, we will

propose a new algorithm allowing a better estimation of the parameter in a reasonable computation time.

**Keywords** Expectation Maximization · Curved exponential · Mixed effect models · Convergence analysis

**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).

## 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:

$$g(y, \theta) = \int_{\mathcal{Z}} f(y, z, \theta) dz.$$

---

Vianney Debavelaere  
Centre de Mathématiques Appliquées, École Polytechnique,  
Palaiseau, France  
E-mail: vianney.debavelaere@polytechnique.edu

Stéphanie Allasonnière  
Centre de Recherche des Cordeliers, Université Paris  
Descartes, Paris, France

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 Dempster et al. (1977); Wu (1983)), 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 Dempster et al. (1977). Its properties have then been studied in numerous papers, see Balakrishnan et al. (2017); Chrétien and Hero (2008); Ma et al. (2000); Meng et al. (1994); Redner and Walker (1984); Tseng (2004); Wu (1983) 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 (Wei and Tanner (1990)) 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 Delyon et al. (1999) 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 Kuhn and Lavielle (2004) 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 (Allasonnière et al., 2010; Bône et al., 2018; Debavelaere et al., 2020; Schiratti et al., 2015), Independent Component Analysis (Allasonniere et al., 2012) and in many medical problems (see Benzekry et al. (2014); Guedj and Perelson (2011); Lavielle and Mentré (2007); Sissoko et al. (2016) 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:

$$f(y, z, \theta) = \exp(-\Psi(\theta) + \langle S(y, z), \Phi(\theta) \rangle), \quad (1)$$

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 (Kuhn et al., 2020; Lartigue et al., 2020; Panhard and Samson, 2009; Samson et al., 2006).

However, this hypothesis can in fact be a bottleneck in lots of situations. For example, it is not verified for heteroscedastic models (Dubois et al., 2011; Kuhn and Lavielle, 2005) nor with more complex models (Bône et al., 2018; Debavelaere et al., 2020; Lindsten, 2013; Meza et al., 2012; Schiratti et al., 2015; Wang, 2007). 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, Kuhn and Lavielle (2005) 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 (Ajmal et al. (2019); Bône et al. (2018); Debavelaere et al. (2020); Schiratti et al. (2015) 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.

## 1 Presentation of the Stochastic Approximation Expectation Maximization (SAEM) Algorithm

In this section, we recall the Stochastic Approximation Expectation Maximization (SAEM) algorithm, first presented in Delyon et al. (1999) 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  $n, l, p \in \mathbb{N}^*$ .

### 1.1 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  $\theta_{k+1} = \operatorname{argmax} Q_k(\theta)$ .

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

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

### 1.2 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  $Q_k(\theta) = Q_{k-1}(\theta) + \gamma_k (\log f(y, z_k, \theta) - Q_{k-1}(\theta))$ . (2)
- *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:

$$f(y, z, \theta) = \exp(-\Psi(\theta) + \langle S(y, z), \Phi(\theta) \rangle), \quad (3)$$

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  $\theta \in \Theta$ ,  $\int_{\mathbb{R}^t} \|S(y, z)\|^2 p(y, z, \theta) \mu(dz) < \infty$  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) 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 1* The assumption (A) can be relaxed by projecting the sequence  $(S_k)_{k \in \mathbb{N}}$  on increasing compacts. See Andrieu et al. (2005) for more details.

Under the hypotheses (M1)-(M5), (SAEM1)-(SAEM4) and (A), it was shown in Andrieu et al. (2005) 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 (Dubois et al., 2011; Kuhn and Lavielle, 2005), 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.

### 1.3 Exponentialization process

We now denote by  $(\theta, \psi)$  the parameters of  $g$  where  $\theta \in \Theta$ ,  $\psi \in \Psi = \mathbb{R}^m$ , and we tackle the case where the model cannot be written under the curved exponential form (4) because of the parameter  $\psi$ . In that case, the log-likelihood can only be written as:

$$f(z, \theta, \psi) = \exp(-\Psi(\theta) + \langle S(z), \Phi(\theta) \rangle) h(z, \psi) \quad (4)$$

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 Kuhn and Lavielle (2005) is to consider  $\psi$  as a Gaussian random variable  $\psi \sim \otimes \mathcal{N}(\bar{\psi}, \sigma^2)$ , where the notation  $\otimes \mathcal{N}(\cdot, \cdot)$  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 section 3).

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

$$\begin{aligned} \log f_\sigma(y, \hat{z}, \hat{\theta}) &= -\Psi(\theta) + \langle S(z), \Phi(\theta) \rangle \\ &+ \log(h(z, \psi)) - \frac{\|\psi - \bar{\psi}\|^2}{2\sigma^2}. \end{aligned} \quad (5)$$

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:

$$\bar{\psi}_{k+1} = \bar{\psi}_k + \gamma_k(\psi_k - \bar{\psi}_k). \quad (6)$$

If we suppose that this augmented model satisfies the hypotheses (M1)-(M5), (SAEM1)-(SAEM4) and (A), we know, using the theorem proved in Andrieu et al. (2005), that it will converge towards a critical point of its incomplete likelihood. However, if this process is used in several applications (Lavielle, 2014), 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 (Kuhn and Lavielle, 2004) as well as for the Approximate SAEM (Allasonnière and Chevallier, 2019).

## 2 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.

### 2.1 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 1** *Assume that the exponentialized model with variance  $\sigma$  verifies the hypotheses (M1)-(M5), (SAEM1)-(SAEM4), (A) and that  $\Psi = \mathbb{R}^m$ . Assume also that, for all  $\sigma > 0$ ,*

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

*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:  $\forall 1 \leq k \leq m$ ,*

$$\int_{\mathbb{R}^m} v_k g(\theta_\infty, \bar{\psi}_\sigma + v) \exp\left(-\frac{\|v\|^2}{2\sigma^2}\right) dv = 0, \quad (7)$$

*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 2* Here, we suppose  $\Psi = \mathbb{R}^m$  to be able to define a Gaussian distribution on  $\Psi$ . 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  $f_\sigma(y, z, \psi + \bar{\psi}_k \mid \theta_k, \bar{\psi}_k)$ .

Under the hypotheses explained section 1.2, it has been shown that this Robbin Monro approximation verifies  $\lim_{k \rightarrow \infty} d((\theta_k, \bar{\psi}_k), \mathcal{L}_\sigma) = 0$ . Moreover, because  $\mathcal{L}_\sigma$  is finite, Delyon et al. (1999) show that the sequence  $(\theta_k, \bar{\psi}_k)_{k \geq 0}$  converges almost surely towards a point  $(\theta_\infty, \bar{\psi}_\sigma) \in \Theta \times \Psi$  (theorem 6). Using the regularity of  $l_\sigma$ , we deduce that those parameters verify  $\partial_{\theta, \bar{\psi}} l_\sigma(\theta_\infty, \bar{\psi}_\sigma) = 0$ .

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

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

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

**Proposition 1** *Suppose that a point  $\psi \in \mathbb{R}^m$  verifies:*

$$\forall v, \theta \in \mathbb{R}^m \times \Theta, g(\theta, \psi + v) = g(\theta, \psi - v). \quad (8)$$

*Then,  $\psi$  is solution to Eq. (7) and can be the parameter returned by the exponential model.*

*Remark 3* It is not necessarily the only possibility of returned parameter. Several values of  $\psi$  could be solutions of Eq. (7).

In particular, a solution of (8) 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 Eq. (7). It is in fact easy to find cases where the maximum is not a solution to Eq. (7) 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 1.

## 2.2 Heuristics

We want to compare the solution of equation (7) to a maximum of the function  $g$ . Because of the form of  $f$  supposed in equation (4), 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 equation (4)).

To explain equation (7), 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 look at the integral:

$$\int_{\Psi} v g(\psi + v) \exp\left(-\frac{v^2}{2\sigma^2}\right) dv$$

for different values of  $\psi$  and  $\sigma$ .  $\psi$  is a solution of equation (7) if and only if this integral is null. It is interesting to remark that one can consider this integral as an expectation if normalized.

First we look at the case  $\psi = \sqrt{2}$ , the  $\operatorname{argmax}$  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 (7) as the integral is strictly positive.

We then reduce the value of  $\sigma$  by taking  $\sigma = 0.1$  on Figure (2b). The function  $v \mapsto g(\sqrt{2} + v) \exp\left(-\frac{v^2}{2\sigma^2}\right)$  is still not symmetric around 0. Even if the value of the integral is smaller,  $\sqrt{2}$  is still not a solution of equation (7).

We now interest ourselves in the case where  $\psi$  is not the  $\operatorname{argmax}$  of  $g$  by taking  $\psi = 1$  and  $\sigma = 1$  on Figure 3. This time,  $g$  is strictly increasing from 1 to  $\sqrt{2}$ . Hence, even if we multiply by the exponential,  $v \mapsto g(1+v) \exp\left(-\frac{v^2}{2\sigma^2}\right)$  is still increasing at 0. As  $g$  decreases slower than it increases, 1 cannot be a solution of equation (7). The integral is indeed strictly positive. The same behaviour would be observed for any point before  $\sqrt{2}$ .

We now look at a value bigger than the maximum:  $\psi = 4$  and  $\sigma = 5$  Figure 4a. This time, as  $g$  decreases at  $v = 4$ ,  $v \mapsto g(4 + v) \exp\left(-\frac{v^2}{2\sigma^2}\right)$  still decreases at 0. But  $g$  decreases slower than it increases. Hence, it is possible to compensate this difference of variation by taking  $\psi > \sqrt{2}$  as a solution of equation (7).

Let us now take a smaller value of  $\sigma$  as in Figure 4b. This time, the integral is negative. Indeed,  $v \mapsto g(4 + v) \exp\left(-\frac{v^2}{2\sigma^2}\right)$  still decreases at 0. However, due to the multiplication by the exponential, the difference of variation before and after the maximum is now way smaller. In particular, it is this time too small to compensate the decrease at 0 and the integral will be negative. To have a solution of equation (7) for this value of  $\sigma$ , we would need to choose a value of  $\psi$  smaller. This suggests that, as  $\sigma$  goes to 0, the solution of equation (7) is closer to  $\sqrt{2}$ .

From these examples, we can deduce two things. First, the  $\operatorname{argmax}$  of  $g$  is not always solution of the equation (7) even for small values of  $\sigma$  because there is a difference in the speed of variation before and after this maximum. However, it is possible to compensate this difference of variation by choosing a value different than the  $\operatorname{argmax}$  and thus to find a solution of (7) different than the  $\operatorname{argmax}$ . Moreover, when  $\sigma$  goes to 0, the difference of variation obtained by multiplying by the exponential is smaller and smaller. It means that a parameter closer and closer to the  $\operatorname{argmax}$  will be solution of equation (7).

We illustrate this behaviour by plotting the exact value of the solution of equation (7) as a function of  $\sigma$  in Figure 5.

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  $\Psi = \mathbb{R}$ . A more general proof in  $\mathbb{R}^m$  for  $m \geq 2$  is given in the supplementary material.

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

**Theorem 2** 1. *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  $\mathcal{L} := \{\psi \in \mathbb{R}^m | \partial_\psi g(\theta, \psi) = 0\}$  is compact and that there exists  $K$  compact such that,  $\forall \sigma > 0$ ,  $\bar{\psi}_\sigma \in K$ . Then,*

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

2. *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*

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

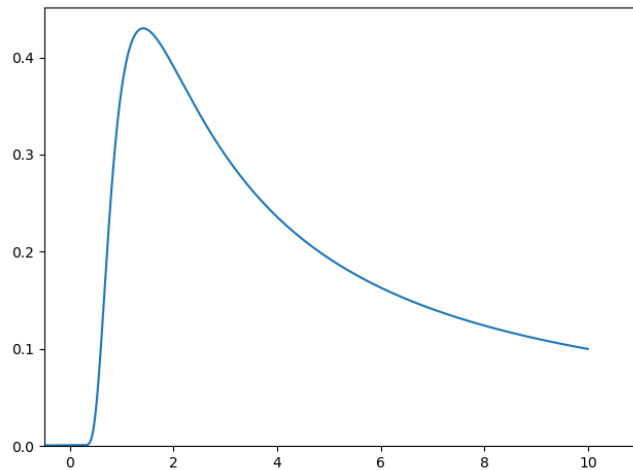
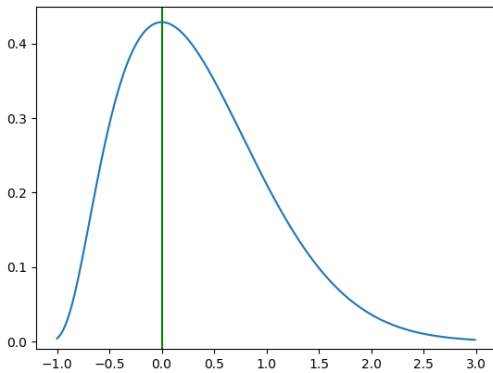
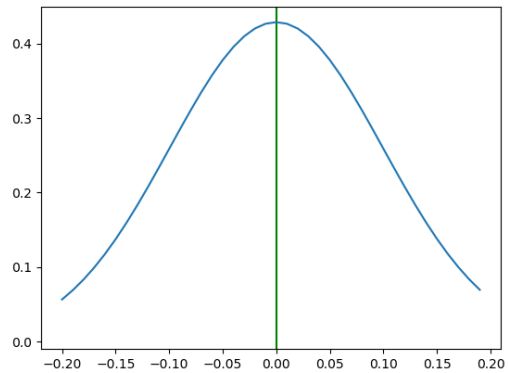


Fig. 1: Function  $g$  studied subsection 2.2, with a maximum reached at  $\sqrt{2}$ .



(a)  $v \mapsto g(\theta, \sqrt{2} + v) \exp\left(\frac{-v^2}{2}\right)$



(b)  $v \mapsto g(\theta, \sqrt{2} + v) \exp\left(\frac{-v^2}{2*0.1^2}\right)$

Fig. 2: Plot of  $v \mapsto g(\theta, \sqrt{2} + v) \exp\left(\frac{-v^2}{2\sigma^2}\right)$  for different values of  $\sigma$ . In all cases, we can see that  $\sqrt{2}$  is not a solution of Eq. (7).

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

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

$$d(\bar{\psi}_\sigma, \mathcal{L}) \leq c\sigma^{\frac{2}{l+2}}. \quad (9)$$

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  $1 \leq k \leq m$ ,

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

*Remark 4* 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 (9) 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 5* 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 6* For  $m = 1$ , we have the exact value of the constant in (9):

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



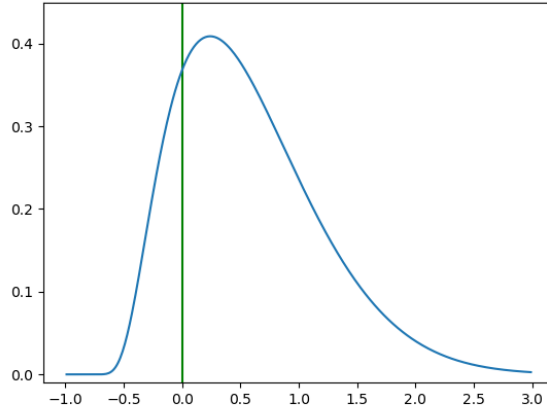
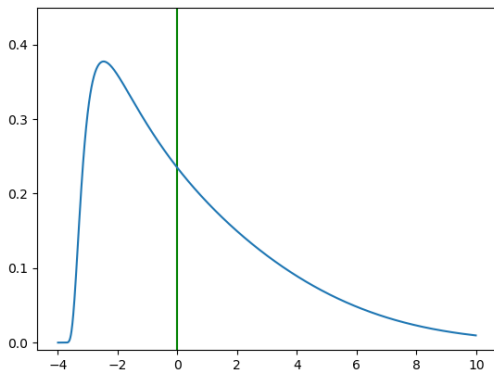
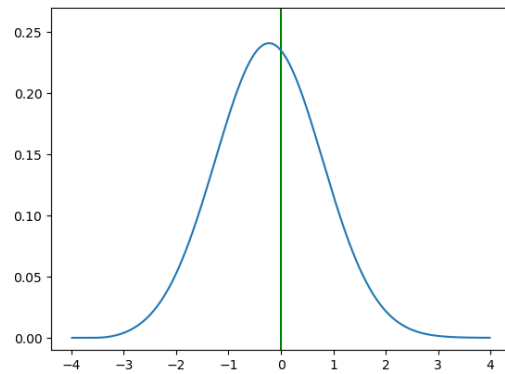


Fig. 3: Plot of  $v \mapsto g(\theta, 1 + v) \exp(-\frac{v^2}{2\sigma^2})$ . Since  $g$  is increasing at 1 and increases quicker than it decreases, 1 is not solution of (7).



(a)  $v \mapsto g(\theta, 4 + v) \exp(-\frac{v^2}{2*5^2})$



(b)  $v \mapsto g(\theta, 4 + v) \exp(-\frac{v^2}{2})$

Fig. 4: Plot of  $v \mapsto g(\theta, 4 + v) \exp(-\frac{v^2}{2\sigma^2})$  for different values of  $\sigma$ .

*Remark 7* The results presented here would still be true in the case of the SAEM-MCMC algorithm. Indeed, the equation (7) verified by the limit parameter of the SAEM would still be verified in the case of the SAEM-MCMC and the same reasoning could then be done.

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 supplementary material.

*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 (7) writes:

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

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, 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  $[\bar{\psi}_\sigma -$

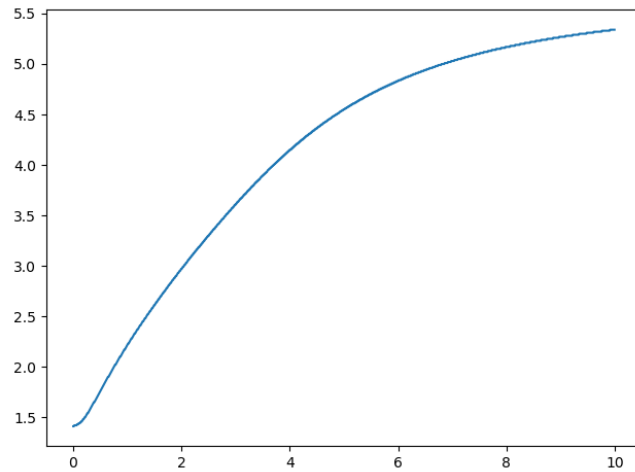


Fig. 5: Solution of Eq. (7) as a function of  $\sigma$  for the function  $g$  studied subsection 2.2.

$c, \bar{\psi}_\sigma + c]$ . 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 (7), 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,

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

On the other hand, we have:

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

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  $g(\bar{\psi}_\sigma + v) - g(\bar{\psi}_\sigma - v) = 2vg'(\tilde{\psi}_v) \geq 2c_0v$ . Hence, we find:

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

But, using an integration per part and defining

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

we have:

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

Hence, because

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

we have:

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

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, by integrating only on  $\{v \geq c\}$ , we first show that

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

Then, by considering this time  $c_1 := \sup\{g'(y) \mid y \in K_0, g'(y) \leq 0\} < 0$ , we find:

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

Hence, for all  $\sigma > 0$ , there exists  $C := 2 \max(c_0, -c_1) > [\bar{\psi}_\sigma - v, \bar{\psi}_\sigma + v]$  and  $\tilde{\psi}_v^1 \in [\bar{\psi}_\sigma - \alpha_\sigma, \bar{\psi}_\sigma + \alpha_\sigma]$  such that: 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 Delyon et al. (1999) 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,  $\forall v \in [\bar{\psi}_\sigma - \alpha_\sigma, \bar{\psi}_\sigma + \alpha_\sigma]$ ,

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

As before, we will split up the integral (7) 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:

$$\left| \int_{|v| \geq \alpha_\sigma} v g(\bar{\psi}_\sigma + v) \exp\left(-\frac{v^2}{2\sigma^2}\right) dv \right| \leq \sigma^2 \|g\|_\infty \exp\left(-\frac{\alpha_\sigma^2}{2\sigma^2}\right). \quad (10)$$

On  $\{v \mid |v| < \alpha_\sigma\}$ , we use twice the mean value theorem to find, for any  $v \in [0, \alpha_\sigma]$ , there exist  $\tilde{\psi}_v^0 \in$

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

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 (7) 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 (10), 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  $\sigma \leq 1$ ,

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

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:

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

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

We use again the equation (7). For all  $\sigma > 0$ ,

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

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

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

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.

### 3 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.

#### 3.1 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 Kuhn and Lavielle (2005) 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  $k_i$ :

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

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:

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

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

*Remark 8* 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 Kuhn and Lavielle (2004), this process converges towards the same limit as the SAEM algorithm. In particular, it has been proved in Kuhn and Lavielle (2005) 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 2 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  $(\beta_1, \beta_2)$  figure 6a. We also plot its behaviour around the maximum and along the axes  $\beta_1$  and  $\beta_2$  figures 6b and 6c.

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 (7) as a function of  $\sigma$ . The results are presented figure 7.

For  $\sigma \geq 1$ , the results of the simulation follow our theory with the estimated parameters estimated close to the solution of the equation (7). 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 particular, it means that the algorithm will be extremely long to converge and, in practice, will stay near the initial value ( $\beta_1 = 6$ ,  $\beta_2 = 34$  here).

### 3.2 Proposition of a new algorithm

To prevent this phenomenon, we now propose a new process 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.

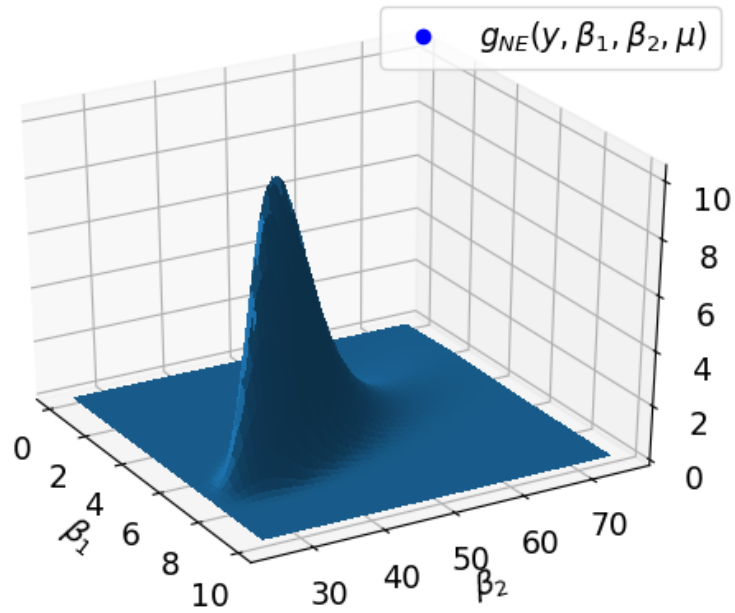
We propose to first run the algorithm with  $\sigma = 1$  for a certain number  $m$  of iterations and then reduce the value of  $\sigma$  by multiplying it by 0.9. We iterate this process every  $m$  iterations until the difference between successive parameters estimated is sufficiently small. We then let the algorithm converge with this small value of  $\sigma$ . This may be seen as launching the algorithm several times with an initialization closer and closer to the true maximum likelihood. While the algorithm will not converge towards the real maximum likelihood estimate as  $\sigma$  is still positive during the last iterations, the error should be smaller than before as  $\sigma$  has been significantly reduced.

To test this new algorithm, we launch this process a hundred times. We present the means and variances of the estimated parameters in Table 1 and as green crosses in figure 7. 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.

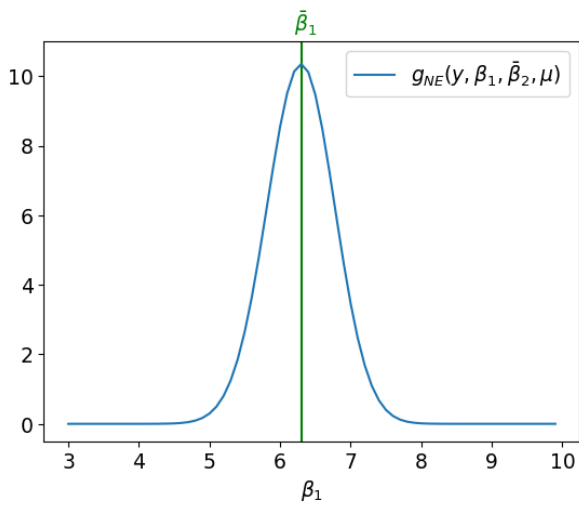
*Remark 9* In Kuhn and Lavielle (2005), the authors use this model and algorithm on a real dataset for different values of  $\sigma$ . They conclude that the estimation of  $(\beta_1, \beta_2)$  does not seem to depend on the choice of  $\sigma$ . In fact, for the particular values of this real dataset, the likelihood is practically symmetric around its maximum. Hence, the error made in that case is indeed small for any  $\sigma$ .

### Conclusion

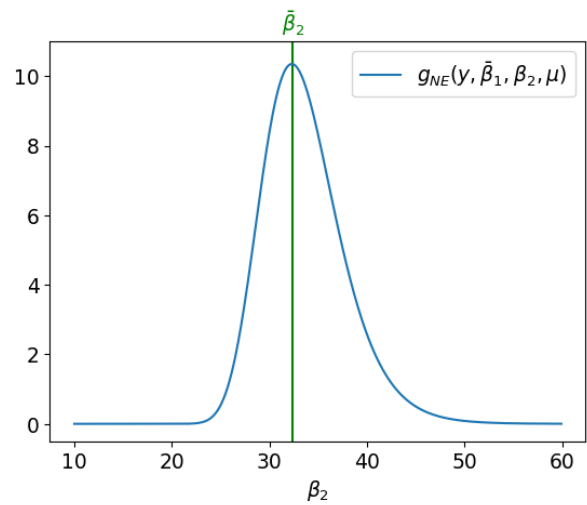
In this paper, we have proved that the exponentialization process does not converge in general towards the maximum likelihood of the initial model using the SAEM or SAEM-MCMC algorithm. If the error converges towards 0 when  $\sigma$  goes to 0, it is numerically impossible to take  $\sigma$  too small 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



(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  $\beta_2$  the argmax of likelihood.



(c) Plot of the incomplete likelihood of the initial model as a function of  $\beta_2$  for  $\beta_1$  the argmax of likelihood.

Fig. 6: Plot of the incomplete likelihood of the initial model as a function of  $(\beta_1, \beta_2)$  along different sections for  $\mu = 5$ .

likelihood does not belong to the curved exponential family.

**References**

Ajmal O, Duchateau L, Kuhn E (2019) Convergent stochastic algorithm for parameter estimation in frailty models using integrated partial likelihood.

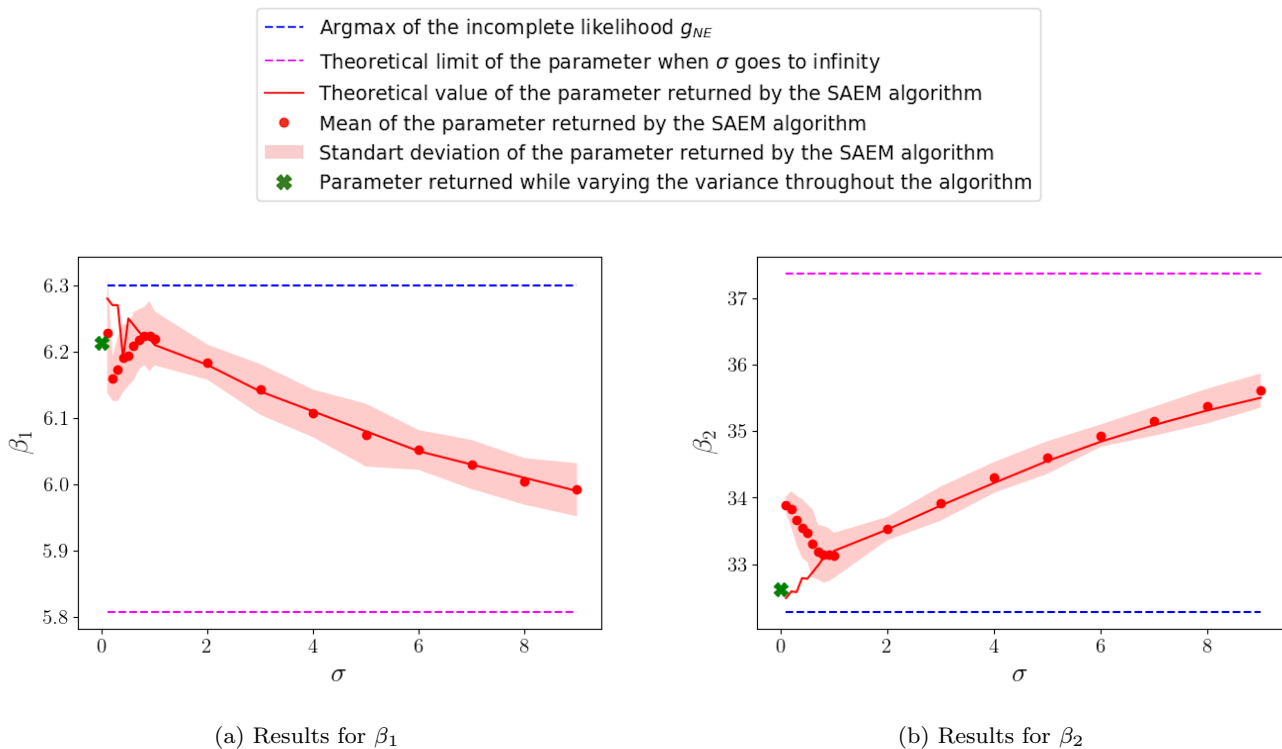


Fig. 7: The red line represents the theoretical value towards which the algorithm is supposed to converge. The red points are the means of the parameters estimated over 100 iterations with their standard deviations represented by the red zone. In dotted blue is the maximum likelihood of the initial model. 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.

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

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  $\beta_2 = 32.28$ .

arXiv preprint arXiv:190907056  
Allasonnière S, Chevallier J (2019) A new class of em algorithms. escaping local minima and handling intractable sampling. preprint  
Allasonnière S, Kuhn E, Trouvé A, et al. (2010) Construction of bayesian deformable models via a stochastic approximation algorithm: a convergence study. *Bernoulli* 16(3):641–678  
Allasonniere S, Younes L, et al. (2012) A stochastic algorithm for probabilistic independent component analysis. *The Annals of Applied Statistics* 6(1):125–160  
Andrieu C, Moulines É, Priouret P (2005) Stability of stochastic approximation under verifiable conditions. *SIAM Journal on control and optimization* 44(1):283–312

Balakrishnan S, Wainwright MJ, Yu B, et al. (2017) Statistical guarantees for the em algorithm: From population to sample-based analysis. *The Annals of Statistics* 45(1):77–120  
Benzekry S, Lamont C, Beheshti A, Tracz A, Ebos JM, Hlatky L, Hahnfeldt P (2014) Classical mathematical models for description and prediction of experimental tumor growth. *PLoS Comput Biol* 10(8):e1003800  
Bône A, Colliot O, Durrleman S (2018) 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*, pp 9271–9280  
Chrétien S, Hero AO (2008) On em algorithms and their proximal generalizations. *ESAIM: Probability and Statistics* 12:308–326

- Debavelaere V, Durrleman S, Allasonnière S (2020) Learning the clustering of longitudinal shape data sets into a mixture of independent or branching trajectories. *International Journal of Computer Vision* pp 1–16
- Delyon B, Lavielle M, Moulines E (1999) Convergence of a stochastic approximation version of the em algorithm. *Annals of statistics* pp 94–128
- Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)* 39(1):1–22
- Dubois A, Lavielle M, Gsteiger S, Pigeolet E, Mentré F (2011) Model-based analyses of bioequivalence crossover trials using the stochastic approximation expectation maximisation algorithm. *Statistics in medicine* 30(21):2582–2600
- Guedj J, Perelson AS (2011) Second-phase hepatitis c virus rna decline during telaprevir-based therapy increases with drug effectiveness: implications for treatment duration. *Hepatology* 53(6):1801–1808
- Kuhn E, Lavielle M (2004) Coupling a stochastic approximation version of em with an mcmc procedure. *ESAIM: Probability and Statistics* 8:115–131
- Kuhn E, Lavielle M (2005) Maximum likelihood estimation in nonlinear mixed effects models. *Computational statistics & data analysis* 49(4):1020–1038
- Kuhn E, Matias C, Rebafka T (2020) Properties of the stochastic approximation em algorithm with mini-batch sampling. *Statistics and Computing* 30(6):1725–1739
- Lartigue T, Durrleman S, Allasonnière S (2020) Deterministic approximate em algorithm; application to the riemann approximation em and the tempered em. *arXiv preprint arXiv:200310126*
- Lavielle M (2014) *Mixed effects models for the population approach: models, tasks, methods and tools*. CRC press
- Lavielle M, Mentré F (2007) Estimation of population pharmacokinetic parameters of saquinavir in hiv patients with the monolix software. *Journal of pharmacokinetics and pharmacodynamics* 34(2):229–249
- Lindsten F (2013) An efficient stochastic approximation em algorithm using conditional particle filters. In: *2013 IEEE International Conference on Acoustics, Speech and Signal Processing, IEEE*, pp 6274–6278
- Ma J, Xu L, Jordan MI (2000) Asymptotic convergence rate of the em algorithm for gaussian mixtures. *Neural Computation* 12(12):2881–2907
- Meng XL, et al. (1994) On the rate of convergence of the em algorithm. *The Annals of Statistics* 22(1):326–339
- Meza C, Osorio F, De la Cruz R (2012) Estimation in nonlinear mixed-effects models using heavy-tailed distributions. *Statistics and Computing* 22(1):121–139
- Panhard X, Samson A (2009) Extension of the saem algorithm for nonlinear mixed models with 2 levels of random effects. *Biostatistics* 10(1):121–135
- Redner RA, Walker HF (1984) Mixture densities, maximum likelihood and the em algorithm. *SIAM review* 26(2):195–239
- Samson A, Lavielle M, Mentré F (2006) Extension of the saem algorithm to left-censored data in nonlinear mixed-effects model: Application to hiv dynamics model. *Computational Statistics & Data Analysis* 51(3):1562–1574
- Schiratti JB, Allasonniere S, Colliot O, Durrleman S (2015) Learning spatiotemporal trajectories from manifold-valued longitudinal data. In: *Advances in neural information processing systems*, pp 2404–2412
- Sissoko D, Laouenan C, Folkesson E, M’lebing AB, Beavogui AH, Baize S, Camara AM, Maes P, Shepherd S, Danel C, et al. (2016) Experimental treatment with favipiravir for ebola virus disease (the jiki trial): a historically controlled, single-arm proof-of-concept trial in guinea. *PLoS medicine* 13(3):e1001967
- Tseng P (2004) An analysis of the em algorithm and entropy-like proximal point methods. *Mathematics of Operations Research* 29(1):27–44
- Wang J (2007) Em algorithms for nonlinear mixed effects models. *Computational statistics & data analysis* 51(6):3244–3256
- Wei GC, Tanner MA (1990) A monte carlo implementation of the em algorithm and the poor man’s data augmentation algorithms. *Journal of the American statistical Association* 85(411):699–704
- Wu CJ (1983) On the convergence properties of the em algorithm. *The Annals of statistics* pp 95–103