A gradient-like Variational Bayesian approach for inverse scattering problems

Leila Gharsalli

A gradient-like Variational Bayesian approach for inverse scattering problems

Leila Gharsalli
Laboratoire des Signaux et Systèmes (L2S)
UMR8506: CNRS-SUPELEC-Univ Paris-Sud
3 rue Joliot-Curie, 91190 Gif-sur-Yvette, France
E-mail: leila.gharsalli@lss.supelec.fr

Abstract

In this document, we present computations of updating shaping parameters for a new method based on the variational Bayesian approach (VBA) allowing to solve a nonlinear inverse scattering problem. The objective is to detect an unknown object from measurements of the scattered field at different frequencies and for several illuminations. This inverse problem is known to be non-linear and ill-posed. So it needs to be regularized by introducing *a priori* information. This is tackled in a Bayesian framework where the particular prior information we account for is that the object is composed of a finite known number of different materials distributed in compact regions. Then we propose the approximate the true joint posterior by a separable law by mean of a gradient-like Variational Bayesian technique. This latter is applied to compute the posterior estimators by allowing a joint update of the shape parameters of the approximating marginals and reconstruct the sought object. The main work is given in [4], while technical details of the variational calculations are presented in the current paper.

1 Introduction

The gradient-like Variational Bayesian approach from now on denoted as GVBA is derived from the classical variational Bayesian approach (VBA, [1]) that aims at approximating a joint posterior distribution $p(\boldsymbol{x}|\boldsymbol{y})$ by a separable law $q(\boldsymbol{x}) = \prod_i q_i(\boldsymbol{x}_i)$ which is as close to the posterior distribution as possible in terms of the Kullback-Leibler divergence. It can be noted that minimizing the KL divergence is equivalent to maximizing the free negative energy derived from statistical physics $\mathcal{F}(\boldsymbol{q}) = \int q(\boldsymbol{x}) \ln\left(p(\boldsymbol{y}, \boldsymbol{x})/q(\boldsymbol{x})\right) \mathrm{d}\boldsymbol{x}$. The solution of this problem can be obtained by alternate optimization with respect to each $q_i(x_i)$ and is given by:

$$q_i(\boldsymbol{x}_i) \propto \exp\left\{\Big\langle \ln(p(\boldsymbol{x}, \boldsymbol{y}))\Big\rangle_{\prod_{j\neq i} q_j(\boldsymbol{x}_j)}\right\}. \tag{1}$$

The computation of $\boldsymbol{q}_i$ requires the knowledge of all $\boldsymbol{q}_j$, $j \neq i$. However, recently, other ways than this classical alternate optimization have been investigated [5]. In fact, the optimization involved in VBA is an infinite dimensional concave problem. Hence, approximating densities $q_i(x_i)$ can be obtained by adapting a classical optimization algorithm, such as a gradient method, to VBA. Using the notion of optimal step, the approximating marginals have an iterative functional form. At iteration n, they read:

$$\begin{aligned} \widetilde{q}_i^{(n)}(\boldsymbol{x}_i) \quad &\propto \quad \left(\widetilde{q}_i^{(n-1)}(\boldsymbol{x}_i)\right)^{(1-\alpha)} \\ &\times \quad \exp\left\{\alpha\Big\langle \ln(p(\boldsymbol{x}, \boldsymbol{y}))\Big\rangle_{\prod_{i\neq j}\widetilde{q}_j^{(n-1)}(x_j)}\right\} \end{aligned} \tag{2}$$

where $\alpha \geq 0$ is a descent step that minimizes the free negative energy at each iteration.

2 Bayesian computations

Let us recall now all the expressions of priors and likelihoods used in the Bayesian framework in the case of a non-linear inverse scattering problem [3]. Then we give the form of separation and expressions of approximating laws for different parameters (we can check several references [2, 3, 6] for the forward modelling of the problem and its formulation). We have:

$$
\begin{array}{ll}
q(\boldsymbol{y}|\boldsymbol{w}, v_\epsilon) = \mathcal{N}(\boldsymbol{G}^o\boldsymbol{w}, v_\epsilon\mathbf{I}), & q(\boldsymbol{w}|\boldsymbol{\chi}, v_\xi) = \mathcal{N}(\boldsymbol{m}_w, \boldsymbol{V}_w), \\
q(\boldsymbol{\chi}|\boldsymbol{z}, \boldsymbol{m}, \boldsymbol{v}) = \mathcal{N}(\boldsymbol{m}_\chi, \boldsymbol{V}_\chi), & q(\boldsymbol{z}) = \exp\left\{\lambda \sum_{\boldsymbol{r}}\sum_{\boldsymbol{r}'} \delta(z(\boldsymbol{r}) - z(\boldsymbol{r}'))\right\}, \\
p(m_k) = \mathcal{N}(m_k|\mu_0, \tau_0), & p(v_k) = \mathcal{IG}(v_k|\eta_0, \phi_0), \\
p(v_\epsilon) = \mathcal{IG}(v_\epsilon|\eta_\epsilon, \phi_\epsilon), & p(v_\xi) = \mathcal{IG}(v_\xi|\eta_\xi, \phi_\xi),
\end{array}
\tag{3}
$$

We denote $\boldsymbol{\psi} = \{\boldsymbol{m}, \boldsymbol{v}, v_\epsilon, v_\xi\}$ the set of the hyper-parameters of the model. Consequently, the joint distribution of the unknowns reads:

$$
\begin{aligned}
p\left(\boldsymbol{\chi}, \boldsymbol{w}, \boldsymbol{z}, \boldsymbol{\psi}|\boldsymbol{y}\right) \;\propto\;& p\left(\boldsymbol{y}|\boldsymbol{w}, v_\epsilon\right) p\left(\boldsymbol{w}|\boldsymbol{\chi}, v_\xi\right) p\left(\boldsymbol{\chi}|\boldsymbol{z}, \boldsymbol{m}, \boldsymbol{v}\right) p\left(\boldsymbol{z}|\lambda\right) p\left(\boldsymbol{m}|\mu_0, \tau_0\right) \\
\times\;& p\left(\boldsymbol{v}|\eta_0, \phi_0\right) p\left(v_\epsilon|\eta_\epsilon, \phi_\epsilon\right) p\left(v_\xi|\eta_\xi, \phi_\xi\right) \\
\propto\;& \frac{\eta_0^{K(\phi_0-1)}\eta_\epsilon^{(\phi_\epsilon-1)}\eta_\xi^{K(\phi_\xi-1)}}{A} \exp\left\{-\frac{||\boldsymbol{y} - \boldsymbol{G}^o\boldsymbol{w}||_2^2}{2v_\epsilon}\right\} \\
\times\;& \exp\left\{-\frac{||\boldsymbol{w} - \boldsymbol{X}\boldsymbol{E}^{inc} - \boldsymbol{X}\boldsymbol{G}^c\boldsymbol{w}||_2^2}{2v_\xi}\right\} \\
\times\;& \exp\left\{-\frac{(\boldsymbol{\chi} - \boldsymbol{m}_\chi)^T{\boldsymbol{V}_\chi}^{-1}(\boldsymbol{\chi} - \boldsymbol{m}_\chi)}{2}\right\} \\
\times\;& \exp\left\{\boldsymbol{\lambda}\sum_{\boldsymbol{i}}\sum_{\boldsymbol{j}} \delta(z(\boldsymbol{i}) - z(\boldsymbol{j})) - T(\boldsymbol{\lambda})\right\} v_\epsilon^{-\phi_\epsilon-1} \\
\times\;& \exp\left\{-\frac{\eta_\epsilon}{v_\epsilon}\right\} v_\epsilon^{-\phi_\xi-1} \exp\left\{-\frac{\eta_\xi}{v_\xi}\right\} \prod_{k=1}^{K} v_k^{-\phi_0-1} \\
\times\;& \exp\left\{-\frac{\eta_0}{v_k}\right\} \exp\left\{-\frac{|m_k - \mu_0|^2}{2\tau_0}\right\} \\
\propto\;& \exp\{\mathcal{L}\},
\end{aligned}
\tag{4}
$$

where $A = (2\pi)^{\frac{M+N(N_P+1)+K}{2}}(v_\epsilon)^{\frac{M}{2}}(v_\xi)^{\frac{N\times N_P}{2}}|\boldsymbol{V}_\chi|^{(-1/2)}(\tau_0)^{(K/2)}\Gamma(\phi_0)^K\Gamma(\phi_\epsilon)\Gamma(\phi_\xi)$ where M is the number of sources, N the number of the elementary square pixels, N_P the number of polarization ($N_P = $ 1D, 2D or 3D) and $\mathcal{L}$ reads:

$$
\begin{aligned}
\mathcal{L} \;=\;& -\left(\frac{M}{2}\right)\log(v_\epsilon) - \left(\frac{NN_P}{2}\right)\log(v_\xi) \\
-\;& \frac{1}{2v_\epsilon}||\boldsymbol{y} - \boldsymbol{G}^o\boldsymbol{w}||_2^2 - \frac{1}{2v_\xi}||\boldsymbol{w} - \boldsymbol{X}\boldsymbol{E}^{inc} - \boldsymbol{X}\boldsymbol{G}^c\boldsymbol{w}||_2^2 \\
-\;& \frac{\sum_k N_k\log(v_k)}{2} - \sum_k\sum_{\boldsymbol{r}}\frac{|\chi(\boldsymbol{r}) - m_k(\boldsymbol{r})|^2}{2v_k} \\
+\;& \boldsymbol{\lambda}\sum_{\boldsymbol{r}}\sum_{\boldsymbol{r}'}\delta(z(\boldsymbol{r}) - z(\boldsymbol{r}')) - \frac{\sum_k|m_k - \mu_0|^2}{2\tau_0} \\
-\;& \sum_k\left(\frac{\eta_0}{v_k} + (\phi_0 + 1)\log(v_k)\right) - \frac{\eta_\epsilon}{v_\epsilon} - (\phi_\epsilon + 1)\log(v_\epsilon) \\
-\;& \frac{\eta_\xi}{v_\xi} - (\phi_\xi + 1)\log(v_\xi).
\end{aligned}
\tag{5}
$$

We may note that applying the joint maximum a posteriori (JMAP) or the posterior mean (PM) to compute the joint posterior distribution (equation (4)) yields intractable form and an approximation is needed to obtain a practical solution. This is done by means of the GVBA. First, a strong separation is chosen:

$$q(\boldsymbol{x}) = q(v_\epsilon)q(v_\xi) \times \prod_i q(\chi_i)q(w_i)q(z_i) \prod_k q(m_k)q(v_k). \tag{6}$$

Then, using equation (2), the approximating marginal for each unknown variable can be computed by means of functional optimization. Updating the approximate posterior requires 7 different gradient steps that we denote by α_w, α_χ, α_z, α_{v_ϵ},α_{v_ξ}, α_{v_k} and α_{m_k}.

2.1 Update shaping parameters

At this point, it can be mentioned that taking the likelihood in the exponential family and using conjugate priors will result in joint posteriors and marginals ranging in the exponential family. Optimization with respect to $q_i(x_i)$ then results in optimizing the parameters of these laws. In the following we summarize the results for the Gaussian - inverse gamma case:

$$\begin{aligned} q(\boldsymbol{w}) &= \mathcal{N}(\widetilde{\boldsymbol{m}}_w, \widetilde{\boldsymbol{V}}_w), & q(\boldsymbol{\chi}) &= \mathcal{N}(\widetilde{\boldsymbol{m}}_\chi, \widetilde{\boldsymbol{V}}_\chi), \\ q(m_k) &= \mathcal{N}(\widetilde{\mu}_k, \widetilde{\tau}_k), & q(v_k) &= \mathcal{IG}(\widetilde{\eta}_k, \widetilde{\phi}_k), \\ q(v_\epsilon) &= \mathcal{IG}(\widetilde{\eta}_\epsilon, \widetilde{\phi}_\epsilon), & q(v_\xi) &= \mathcal{IG}(\widetilde{\eta}_\xi, \widetilde{\phi}_\xi), \end{aligned}$$

$$q(\boldsymbol{z}) = \widetilde{\boldsymbol{\zeta}}_k \propto exp\left(\lambda \sum_{\boldsymbol{r}\in\mathcal{D}} \sum_{\boldsymbol{r}'\in V(\boldsymbol{r})} \widetilde{\zeta}(\boldsymbol{r}')\right), \tag{7}$$

where tilded parameters are given in the following. The key used to obtain these forms is equation (2).

2.1.1 Contrast source $\boldsymbol{w}$

$$\begin{aligned} \log\left(\widetilde{q}^n(w(\boldsymbol{i}))\right) &\propto (1-\alpha_w)\log\left(\widetilde{q}(w(\boldsymbol{i}))\right) + \alpha_w \left\langle \log\left(p\left(\boldsymbol{\chi}, \boldsymbol{w}, \boldsymbol{z}, \boldsymbol{\psi}, \boldsymbol{y}\right)\right)\right\rangle_{\tilde{q}(/w(\boldsymbol{i}))} \\ &\propto (1-\alpha_w)\left[\widetilde{\boldsymbol{v}}_{w(\boldsymbol{i})} w(\boldsymbol{i})^2 - 2\widetilde{\boldsymbol{m}}_{w(\boldsymbol{i})} w(\boldsymbol{i})\right] \\ &+ \alpha_w \left\langle \log\left(p(\boldsymbol{y}|\boldsymbol{w}, v_\epsilon)\right) + \log\left(p(\boldsymbol{w}|\boldsymbol{\chi}, v_\xi)\right)\right\rangle_{\tilde{q}(/w(\boldsymbol{i}))} \\ &\propto (1-\alpha_w)\left[\widetilde{\boldsymbol{v}}_{w(\boldsymbol{i})} w(\boldsymbol{i})^2 - 2\widetilde{\boldsymbol{m}}_{w(\boldsymbol{i})} w(\boldsymbol{i})\right] \\ &- \frac{\alpha_w}{2}\left\langle \frac{||\boldsymbol{y} - \boldsymbol{G}^o\boldsymbol{w}||_2^2}{v_\epsilon} + \frac{||\boldsymbol{w} - \boldsymbol{X}\boldsymbol{E}^{inc} - \boldsymbol{X}\boldsymbol{G}^c\boldsymbol{w}||_2^2}{v_\xi}\right\rangle_{\tilde{q}(/w(\boldsymbol{i}))} \end{aligned} \tag{8}$$

where $q(/w(\boldsymbol{i})) = \prod_{j\neq i} q(w(\boldsymbol{j}))q(\boldsymbol{\chi})q(\boldsymbol{z})q(\boldsymbol{\psi})$. Or

$$\begin{aligned} \left\langle ||\boldsymbol{y} - \boldsymbol{G}^o\boldsymbol{w}||_2^2\right\rangle_{\tilde{q}(/w(\boldsymbol{i}))} &\propto \sum_{\boldsymbol{l}} |G^o(\boldsymbol{l}, \boldsymbol{i})|^2 |w(\boldsymbol{i})|^2 \\ &+ 2\,\Re e\left(\sum_{\boldsymbol{k}} G^{o*}(\boldsymbol{k}, \boldsymbol{i})\left[y(\boldsymbol{k}) - \sum_{\boldsymbol{j}\neq\boldsymbol{i}} G^o(\boldsymbol{k}, \boldsymbol{j})\widetilde{m}_w(\boldsymbol{j})\right] w^*(\boldsymbol{i})\right), \end{aligned}$$

and

$$
\begin{aligned}
\left\langle ||\boldsymbol{w} - \boldsymbol{X}\boldsymbol{E}^{inc} - \boldsymbol{X}\boldsymbol{G}^c\boldsymbol{w}||_2^2 \right\rangle_{\tilde{q}(/w(\boldsymbol{i}))} \propto\ & |w(\boldsymbol{i})|^2 - G^{c*}(\boldsymbol{i},\boldsymbol{i})\ \bar{\chi}^*(\boldsymbol{i})\ |w(\boldsymbol{i})|^2 - G^c(\boldsymbol{i},\boldsymbol{i})\ \bar{\chi}(\boldsymbol{i})\ |w(\boldsymbol{i})|^2 \\
+\ & \sum_{\boldsymbol{j}} |G^c(\boldsymbol{j},\boldsymbol{i})|\ \overline{|\chi(\boldsymbol{i})|^2}\ |w(\boldsymbol{i})|^2 - 2\ \Re e\left(E^{inc}(\boldsymbol{i})\bar{\chi}(\boldsymbol{i})w^*(\boldsymbol{i})\right) \\
+\ & 2\ \Re e\left(\sum_{\boldsymbol{j}} |G^{c*}(\boldsymbol{j},\boldsymbol{i})|\ \overline{|\chi(\boldsymbol{j})|^2}\ |w^*(\boldsymbol{j})|^2\right) \\
-\ & 2\ \Re e\left(\sum_{\boldsymbol{j}\neq\boldsymbol{i}} \bar{\chi}(\boldsymbol{i})\ G^c(\boldsymbol{i},\boldsymbol{j})\ \bar{w}(\boldsymbol{j})\ w^*(\boldsymbol{i})\right) \\
-\ & 2\ \Re e\left(\sum_{\boldsymbol{j}\neq\boldsymbol{i}} \bar{\chi}^*(\boldsymbol{j})\ G^{c*}(\boldsymbol{j},\boldsymbol{i})\ \bar{w}(\boldsymbol{j})\ w^*(\boldsymbol{i})\right) \\
+\ & 2\ \Re e\left(\sum_{\boldsymbol{j}} G^{c*}(\boldsymbol{j},\boldsymbol{i})\ \overline{|\chi(\boldsymbol{j})|^2}\ \sum_{\boldsymbol{k}\neq\boldsymbol{i}} G^c(\boldsymbol{j},\boldsymbol{k})\ \bar{w}(\boldsymbol{k})\ w^*(\boldsymbol{i})\right) \\
+\ & cte.
\end{aligned}
$$

where * denotes the conjugate complex.

Hence, by combining all the terms, the approximating law, at the iteration n, becomes $q^n(\boldsymbol{w}) = \mathcal{N}(\widetilde{\boldsymbol{m}}_w^n, \widetilde{\boldsymbol{V}}_w^n)$ where:

$$
\begin{aligned}
\widetilde{V}_w^n(\boldsymbol{i}) = \Bigg[& (1-\alpha_w)\widetilde{V}_w(\boldsymbol{i}) + \alpha_w \Bigg(\overline{v_\epsilon^{-1}} \sum_{\boldsymbol{j}} |G^o(\boldsymbol{j},\boldsymbol{i})|^2 \\
& + \overline{v_\xi^{-1}} \Big[1 - 2\ \Re e\big(G^c(\boldsymbol{i},\boldsymbol{i})\widetilde{m}_\chi(\boldsymbol{i})\big) + \sum_{\boldsymbol{j}} |G^c(\boldsymbol{j},\boldsymbol{i})|^2 \big(|\widetilde{m}_\chi(\boldsymbol{i})|^2 + \widetilde{V}_\chi(\boldsymbol{i})\big) \Big] \Bigg) \Bigg]^{-1},
\end{aligned}
$$

$$
\begin{aligned}
\frac{\widetilde{m}_w^n(\boldsymbol{i})}{\widetilde{V}_w^n(\boldsymbol{i})} =\ & (1-\alpha_w)\,\widetilde{m}_w(\boldsymbol{i}) + \alpha_w \Bigg[\overline{v_\epsilon^{-1}} \sum_{\boldsymbol{k}} G^{o*}(\boldsymbol{k},\boldsymbol{i}) \Big(y(\boldsymbol{k}) - \sum_{j\neq i} G^o(\boldsymbol{k},\boldsymbol{j})\widetilde{m}_w(\boldsymbol{j}) \Big) \\
+\ & \overline{v_\xi^{-1}} \Bigg(E^{inc}(\boldsymbol{i})\widetilde{m}_\chi(\boldsymbol{i}) - \sum_{\boldsymbol{j}} G^{c*}(\boldsymbol{j},\boldsymbol{i}) \Big(|\widetilde{m}_\chi(\boldsymbol{j})|^2 + \widetilde{V}_\chi(\boldsymbol{j}) \Big) E^{inc}(\boldsymbol{j}) \\
+\ & \sum_{\boldsymbol{j}\neq\boldsymbol{i}} \widetilde{m}_\chi(\boldsymbol{i}) G^c(\boldsymbol{i},\boldsymbol{j})\widetilde{m}_w(\boldsymbol{j}) + \sum_{\boldsymbol{j}\neq\boldsymbol{i}} \widetilde{m}_\chi^*(\boldsymbol{i}) G^{c*}(\boldsymbol{j},\boldsymbol{i})\widetilde{m}_w(\boldsymbol{j}) \\
-\ & \sum_{\boldsymbol{j}} G^{c*}(\boldsymbol{j},\boldsymbol{i}) \Big(|\widetilde{m}_\chi(\boldsymbol{i})|^2 + \widetilde{V}_\chi(\boldsymbol{i}) \Big) \sum_{\boldsymbol{k}\neq\boldsymbol{j}} G^c(\boldsymbol{j},\boldsymbol{k})\overline{w}(\boldsymbol{k}) \Bigg) \Bigg],
\end{aligned}
$$

By adding the missing $\widetilde{m}_w$ to the right-hand side of the previous equation and putting it in a vectorial form, we obtain the final form:

$$
\begin{cases}
\widetilde{\boldsymbol{V}}_w^n = \Big[(1-\alpha_w)\,\widetilde{\boldsymbol{V}}_w^{-1} + \alpha_w\ \mathbf{Diag}\left(\overline{v_\epsilon^{-1}}\Gamma^o + \overline{v_\xi^{-1}}\overline{\Gamma^{xc}} \right) \Big]^{-1} \\
\widetilde{\boldsymbol{m}}_w^n = \widetilde{\boldsymbol{m}}_w + \alpha_w\ \widetilde{\boldsymbol{V}}_w^n \Big[\Big[\overline{v_\epsilon^{-1}} \boldsymbol{G}^{o^H} (\boldsymbol{y} - \boldsymbol{G}^o\widetilde{\boldsymbol{m}}_w) \\
\quad + \overline{v_\xi^{-1}} \Big(\widetilde{\boldsymbol{m}}_\chi \boldsymbol{E}^{inc} - \boldsymbol{G}^{c^H} \left(\widetilde{\boldsymbol{m}}_\chi^2 + \widetilde{\boldsymbol{V}}_\chi \right) \boldsymbol{E}^{inc} \\
\quad - \widetilde{\boldsymbol{m}}_w + \widetilde{\boldsymbol{m}}_\chi \boldsymbol{G}^c \widetilde{\boldsymbol{m}}_w + \boldsymbol{G}^{c^H} \widetilde{\boldsymbol{m}}_\chi^* \widetilde{\boldsymbol{m}}_w - \boldsymbol{G}^{c^H} \left(\widetilde{\boldsymbol{m}}_\chi^2 + \widetilde{\boldsymbol{V}}_\chi \right) \boldsymbol{G}^c \widetilde{\boldsymbol{m}}_w \Big) \Big] \Big]
\end{cases}
\tag{9}
$$

where overline denotes the expectation of the variable with respect to q (i.e. $\overline{u} = \mathrm{E}(u)_q$), superscript ^{H} indicates the conjugate transpose and Γ^o and Γ^{xc} are given as:

$$
\begin{aligned}
\Gamma^o(\boldsymbol{i}) &= \sum_{\boldsymbol{j}} |G^o(\boldsymbol{j},\boldsymbol{i})|^2, \\
\overline{\Gamma^{xc}}(\boldsymbol{i}) &= 1 - 2\Re e(G^c(\boldsymbol{i},\boldsymbol{i})\widetilde{m}_\chi(\boldsymbol{i})) \\
&+ (\widetilde{m}_\chi^2(\boldsymbol{i}) + \widetilde{V}_\chi(\boldsymbol{i}))\sum_{\boldsymbol{j}} |G^c(\boldsymbol{j},\boldsymbol{i})|^2
\end{aligned}
$$

2.1.2 Contrast $\boldsymbol{\chi}$

$$
\begin{aligned}
\log\left(\widetilde{q}^n(\chi(\boldsymbol{i}))\right) &\propto (1-\alpha)\log\left(\widetilde{q}(\chi(\boldsymbol{i}))\right) + \alpha_\chi \left\langle \log\left(p\left(\boldsymbol{\chi},\boldsymbol{w},\boldsymbol{z},\boldsymbol{\psi},\boldsymbol{y}\right)\right)\right\rangle_{\tilde{q}(/\chi(\boldsymbol{i}))} \\
&\propto (1-\alpha_\chi)\left[\widetilde{\boldsymbol{v}}_{\chi(\boldsymbol{i})}\chi(\boldsymbol{i})^2 - 2\widetilde{\boldsymbol{m}}_{\chi(\boldsymbol{i})}\chi(\boldsymbol{i})\right] \\
&+ \alpha_\chi \left\langle \log(p(\boldsymbol{w}|\boldsymbol{\chi},v_\epsilon)) + \log\left(p(\boldsymbol{\chi}|\boldsymbol{z},\boldsymbol{v},\boldsymbol{m})\right)\right\rangle_{\tilde{q}(/\chi(\boldsymbol{i}))} \\
&\propto (1-\alpha_\chi)\left[\widetilde{\boldsymbol{v}}_{\chi(\boldsymbol{i})}\chi(\boldsymbol{i})^2 - 2\widetilde{\boldsymbol{m}}_{\chi(\boldsymbol{i})}\chi(\boldsymbol{i})\right] \\
&+ \alpha_\chi \left\langle \frac{||\boldsymbol{w}-\boldsymbol{X}\boldsymbol{E}||_2^2}{v_\xi} + \left(\boldsymbol{\chi}-\boldsymbol{m}_\chi\right)^T \boldsymbol{V}_\chi^{-1}\left(\boldsymbol{\chi}-\boldsymbol{m}_\chi\right)\right\rangle_{\tilde{q}(/\chi(\boldsymbol{i}))},
\end{aligned} \tag{10}
$$

where $q(/\chi(\boldsymbol{i})) = \prod_{j\neq i} q(\chi(\boldsymbol{j}))q(\boldsymbol{w})q(\boldsymbol{z})q(\boldsymbol{\psi})$, or

$$
\begin{aligned}
\left\langle ||\boldsymbol{w}-\boldsymbol{X}\boldsymbol{E}||_2^2\right\rangle_{\tilde{q}(/\chi(\boldsymbol{i}))} &\propto -2\Re e\left(\overline{w^*(\boldsymbol{i})E(\boldsymbol{i})}\chi(\boldsymbol{i})\right) + \overline{E^2}(\boldsymbol{i})|\chi(\boldsymbol{i})|^2 \\
&+ \left\langle ||\boldsymbol{w}||_2^2 - 2\Re e\left(\sum_{\boldsymbol{j}\neq\boldsymbol{i}} w^*(\boldsymbol{j})E(\boldsymbol{j})\chi(\boldsymbol{j})\right) + \sum_{\boldsymbol{j}\neq\boldsymbol{i}} |E(\boldsymbol{j})\chi(\boldsymbol{j})|^2\right\rangle_{\tilde{q}(/\chi(\boldsymbol{i}))},
\end{aligned}
$$

and

$$
\begin{aligned}
\left\langle \left(\boldsymbol{\chi}-\boldsymbol{m}_\chi\right)^T \boldsymbol{V}_\chi^{-1}\left(\boldsymbol{\chi}-\boldsymbol{m}_\chi\right)\right\rangle_{\tilde{q}(/\chi(\boldsymbol{i}))} &\propto \sum_{\boldsymbol{j}} \left\langle \frac{||\chi(\boldsymbol{j})-\boldsymbol{m}_k||_2^2}{v_k}\right\rangle_{\tilde{q}(/\chi(\boldsymbol{i}))} \\
&\propto \overline{v_k^{-1}}\left(|\chi(\boldsymbol{i})|^2 - 2\Re e\left(\widetilde{\boldsymbol{m}}_k\chi^*(\boldsymbol{i})\right)\right),
\end{aligned}
$$

By identification, we find

$$
\widetilde{V}_\chi^n(\boldsymbol{i}) \propto \left[(1-\alpha_\chi)\widetilde{V}_\chi(\boldsymbol{i})^{-1} + \alpha_\chi\left(\overline{v_\xi^{-1}}\,\overline{E^2(\boldsymbol{i})} + \overline{\widetilde{v}_\chi^{-1}(\boldsymbol{i})}\right)\right]^{-1},
$$

and

$$
\frac{\widetilde{m}_\chi^n(\boldsymbol{i})}{\widetilde{V}_\chi^n(\boldsymbol{i})} \propto (1-\alpha_\chi)\widetilde{m}_\chi(\boldsymbol{i}) + \alpha_\chi\left[\frac{\overline{\widetilde{m}_\chi(\boldsymbol{i})}}{\widetilde{V}_\chi(\boldsymbol{i})} + \overline{v_\xi^{-1}}\,\overline{w(\boldsymbol{i})E^*(\boldsymbol{i})}\right],
$$

where $\overline{\boldsymbol{w}\boldsymbol{E}^*}$ is the mean of the vector $\boldsymbol{w}\boldsymbol{E}^*$ such that:

$$
\begin{aligned}
\overline{wE^*}(\boldsymbol{i}) &= \sum_{N_f N_v N} E^{inc*}(\boldsymbol{i})\widetilde{m}_w(\boldsymbol{i}) + \widetilde{m}_w(\boldsymbol{i})\sum_{jb} G^{c*}(\boldsymbol{i},\boldsymbol{i}')\widetilde{m}_w^*(\boldsymbol{i}') \\
&+ G^{c*}(\boldsymbol{i},\boldsymbol{i})\widetilde{V}_w(\boldsymbol{i}),
\end{aligned}
$$

and $\overline{\boldsymbol{E}^2}$ is a diagonal matrice whose elements are written such that:

$$
\begin{aligned}
\overline{E^2}(\boldsymbol{i}) &= \sum_{N_f N_v N_P} |E^{inc}(\boldsymbol{i})|^2 + 2\Re e\left(E^{inc*}(\boldsymbol{i})G^c\widetilde{m}_w(\boldsymbol{i})\right) \\
&+ |\sum_{\boldsymbol{i}'} G^c(\boldsymbol{i},\boldsymbol{i}')\widetilde{m}_w(\boldsymbol{i})|^2 + \sum_{\boldsymbol{i}'} |G^c(\boldsymbol{i},\boldsymbol{i}')|^2\widetilde{V}_w(\boldsymbol{i}),
\end{aligned}
$$

Then the vectorial form is given by:

$$\left\{\begin{array}{l} \widetilde{\boldsymbol{V}}_{\chi}^{n}=\left[(1-\alpha_{\chi})\,\widetilde{\boldsymbol{V}}_{\chi}^{-1}+\;\alpha_{\chi}\left(\mathbf{Diag}\left(\overline{v_{\xi}^{-1}}\overline{\boldsymbol{E}^{2}}+\overline{\boldsymbol{V}_{\chi}^{-1}}\right)\right)\right]^{-1} \\ \widetilde{\boldsymbol{m}}_{\chi}^{n}=\alpha_{\chi}\,\widetilde{\boldsymbol{V}}_{\chi}^{n}\left[\left(\sum_{k}\overline{v_{k}^{-1}}\widetilde{\boldsymbol{\zeta}}_{k}\circ\widetilde{\boldsymbol{m}}_{k}+\overline{v_{\xi}^{-1}}\overline{\boldsymbol{w}\circ\boldsymbol{E}^{*}}\right)\right] \end{array}\right. \tag{11}$$

where $\overline{V_{\chi}^{-1}}(\boldsymbol{i},\boldsymbol{i})=\sum_{k}\widetilde{\boldsymbol{\zeta}}_{k}(\boldsymbol{i})\overline{v_{k}^{-1}}$.

2.1.3 The hidden field $\boldsymbol{z}$

$$\begin{aligned}
\log\left(\widetilde{q}^{n}(z(\boldsymbol{i}))\right) &\propto (1-\alpha_{z})\log\left(\widetilde{q}(z(\boldsymbol{i}))\right)\;+\;\alpha\;\langle\log\left(p\left(\boldsymbol{\chi},\boldsymbol{w},\boldsymbol{z},\boldsymbol{\psi},\boldsymbol{y}\right)\right)\rangle_{\tilde{q}(/z(\boldsymbol{i}))} \\
&\propto (1-\alpha_{z})\log\left(\widetilde{q}(z(\boldsymbol{i}))\right)\;+\;\alpha_{z}\;\Bigg\langle-\frac{1}{2}\log(v_{k})+\frac{||\chi(\boldsymbol{i})-m_{\chi}(\boldsymbol{i})||_{2}^{2}}{2v_{k}} \\
&- \lambda\sum_{\boldsymbol{l}\in\mathcal{V}(\boldsymbol{i})}\delta\left(z(\boldsymbol{l})-k\right)\Bigg\rangle_{\tilde{q}(/z(\boldsymbol{i}))} \\
&\propto (1-\alpha)\log\left(\widetilde{\zeta}_{k}(\boldsymbol{i})\right)-\frac{\alpha_{z}}{2}\Bigg[\overline{\log(v_{k})}+\overline{v_{k}^{-1}}\;\langle|\chi(\boldsymbol{i})-m_{\chi}(\boldsymbol{i})|\rangle_{\tilde{q}(/z(\boldsymbol{i}))} \\
&- 2\lambda\sum_{\boldsymbol{j}\in\mathcal{V}(\boldsymbol{i})}\widetilde{\zeta}_{k}(\boldsymbol{j})\Bigg] \\
&\propto (1-\alpha_{z})\log\left(\widetilde{\zeta}_{k}(\boldsymbol{i})\right)-\frac{\alpha_{z}}{2}\Bigg[\Psi(\tilde{\eta}_{k})-\log\tilde{\phi}_{k}+\overline{v_{k}^{-1}}\Big[|\widetilde{m}_{\chi}(\boldsymbol{i})|^{2} \\
&+ \widetilde{m}_{k}^{\dagger}-2\Re e\left(\widetilde{m}_{\chi}^{\dagger}\widetilde{m}_{\chi}^{*}(\boldsymbol{i})\right)\Big]+\lambda\sum_{\boldsymbol{j}\in\mathcal{V}(\boldsymbol{i})}\widetilde{\zeta}_{k}(\boldsymbol{j})\Bigg],
\end{aligned}$$

where $q(/z(\boldsymbol{i}))=\prod_{j\neq i}q(z(\boldsymbol{j}))q(\boldsymbol{\chi})q(\boldsymbol{w})q(\boldsymbol{\psi})$ and $\widetilde{m}_{\chi}^{\dagger 2}=\widetilde{m}_{k}^{2}+\widetilde{\tau}_{k}$.

By identification, it becomes:

$$\begin{aligned}
\widetilde{\zeta}_{k}^{n} &= \widetilde{\zeta}_{k}^{(1-\alpha_{z})}\exp\Bigg\{-\frac{\alpha_{z}}{2}\Bigg(\Psi(\widetilde{\eta}_{k})+\log(\widetilde{\phi}_{k})+\overline{v_{k}^{-1}}\Big((\widetilde{m}_{\chi}(\boldsymbol{i})-\widetilde{\mu}_{k})^{2}+\widetilde{\tau}_{k} \\
&+ \widetilde{V}_{\chi}(\boldsymbol{r})\Big)-\lambda\sum_{\boldsymbol{i}'\in\mathcal{V}(\boldsymbol{r})}\widetilde{\boldsymbol{\zeta}}_{k}(\boldsymbol{i}')\Bigg)\Bigg\},
\end{aligned} \tag{12}$$

where $\Psi(x)=\frac{\partial}{\partial x}\log\Gamma(x)$ the digamma function with $\Gamma(x)$ is the gamma function.

2.1.4 The observation variance v_{ϵ}

$$\begin{aligned}
\log\left(\widetilde{q}^{n}(v_{\epsilon})\right) &\propto (1-\alpha_{v_{\epsilon}})\log\left(\widetilde{q}(v_{\epsilon})\right)\;+\;\alpha_{v_{\epsilon}}\;\langle\log\left(p\left(\boldsymbol{\chi},\boldsymbol{w},\boldsymbol{z},\boldsymbol{\psi},\boldsymbol{y}\right)\right)\rangle_{\tilde{q}(/v_{\epsilon})} \\
&\propto (1-\alpha_{v_{\epsilon}})\left(-\frac{v_{\epsilon}}{\tilde{\eta}_{\epsilon}}+(\tilde{\phi}_{\epsilon}\log(v_{\epsilon}))\right) \\
&+ \alpha_{v_{\epsilon}}\;\langle p(\boldsymbol{y}|\boldsymbol{w},v_{\epsilon})p(v_{\epsilon}|\eta_{\epsilon},\phi_{\epsilon})\rangle_{\tilde{q}(/v_{\epsilon})}\,,
\end{aligned} \tag{13}$$

where $q(/v_{\epsilon}))=q(\boldsymbol{\chi})q(\boldsymbol{w}))q(\boldsymbol{z})\prod_{l\neq v_{\epsilon}}q(\psi_{l})$, with:

$$\begin{aligned}
\langle p(\boldsymbol{y}|\boldsymbol{w},v_{\epsilon})p(v_{\epsilon}|\eta_{\epsilon},\phi_{\epsilon})\rangle_{\tilde{q}(/v_{\epsilon})} &\propto \left\langle-\frac{M}{2}\log(v_{\epsilon})-\frac{||\boldsymbol{y}-\boldsymbol{G}^{o}\boldsymbol{w}||_{2}^{2}}{2v_{\epsilon}}-\frac{\eta_{\epsilon}}{v_{\epsilon}}-(\phi_{\epsilon}-1)\log(v_{\epsilon})\right\rangle_{\tilde{q}(/v_{\epsilon})} \\
&\propto -(1+\phi_{\epsilon}+\frac{M}{2})\log(v_{\epsilon})-v_{\epsilon}^{-1}\left\langle\eta_{\epsilon}+\frac{||\boldsymbol{y}-\boldsymbol{G}^{o}\boldsymbol{w}||_{2}^{2}}{2}\right\rangle_{\tilde{q}(/v_{\epsilon})},
\end{aligned}$$

where

$$
\begin{aligned}
\langle ||\boldsymbol{y}-\boldsymbol{G}^o\boldsymbol{w}||_2^2\rangle_{\tilde{q}(/v_\epsilon)} &\propto \sum_{\boldsymbol{i}}\Bigg(|y(\boldsymbol{i})|^2+\sum_{\boldsymbol{j}}\sum_{\boldsymbol{k}}G^o(\boldsymbol{i},\boldsymbol{k})\widetilde{m}_w(\boldsymbol{j})\widetilde{m}_w^*(\boldsymbol{k})+\sum_{\boldsymbol{j}}|G^o(\boldsymbol{i},\boldsymbol{j})|\widetilde{v}_w(\boldsymbol{j}) \\
&- 2\Re e\Big(y^*\sum_{\boldsymbol{j}}G^o(\boldsymbol{i},\boldsymbol{j})\widetilde{m}_w^*(\boldsymbol{j})\Big)\Bigg),
\end{aligned}
$$

By identification, we obtain:

$$
\begin{cases}
\widetilde{\phi}_\epsilon^n = (1-\alpha_{v_\epsilon})\,\widetilde{\phi}_\epsilon+\alpha_{v_\epsilon}\left(\phi_\epsilon+\frac{M}{2}\right) \\
\widetilde{\eta}_\epsilon^n = \frac{1-\alpha_{v_\epsilon}}{\widetilde{\eta}_\epsilon}+\alpha_{v_\epsilon}\left(\eta_\epsilon+\frac{1}{2}\Big(||\boldsymbol{y}||_2^2+||\boldsymbol{G}^o\widetilde{\boldsymbol{m}}_w||_2^2-2\Re e(\boldsymbol{y}^H\boldsymbol{G}^o\widetilde{\boldsymbol{m}}_w)+||\boldsymbol{G}^{o2}\widetilde{\boldsymbol{v}}_w||_1\Big)\right)
\end{cases} \tag{14}
$$

2.1.5 Coupling noise v_ξ

$$
\begin{aligned}
\log\left(\widetilde{q}^n(v_\xi)\right) &\propto \left(1-\alpha_{v_\xi}\right)\log\left(\widetilde{q}(v_\xi)\right)\;+\;\alpha_\xi\;\langle\log\left(p\left(\boldsymbol{\chi},\boldsymbol{w},\boldsymbol{z},\boldsymbol{\psi},\boldsymbol{y}\right)\right)\rangle_{\tilde{q}(/v_\xi)} \\
&\propto \left(1-\alpha_{v_\xi}\right)\left(-\frac{v_\xi}{\tilde{\eta}_\xi}+(\tilde{\phi}_\xi-\log(v_\xi))\right) \\
&+ \alpha_{v_\xi}\;\langle p(\boldsymbol{w}|\boldsymbol{\chi},v_\xi)p(v_\xi|\eta_\xi,\phi_\xi)\rangle_{\tilde{q}(/v_\xi)}\,,
\end{aligned} \tag{15}
$$

where $q(/v_\xi))=q(\boldsymbol{\chi})q(\boldsymbol{w}))q(\boldsymbol{z})\prod_{l\neq v_\xi}q(\psi_l)$ with

$$
\begin{aligned}
\langle p(\boldsymbol{w}|\boldsymbol{\chi},v_\xi)p(v_\xi|\eta_\xi,\phi_\xi)\rangle_{\tilde{q}(/v_\xi)} &\propto \left\langle\log v_\xi^{-\frac{N_PN}{2}}-\frac{||\boldsymbol{w}-\boldsymbol{X}\boldsymbol{E}||_2^2}{2v_\xi}-\frac{\eta_\xi}{v_\xi}-(\phi_\xi+1)\log(v_\xi)\right\rangle_{\tilde{q}(/v_\xi)} \\
&\propto \left(\eta_\xi+\frac{N_PN}{2}\right)\log\left(v_\xi\right)-v_\xi^{-1}\left\langle\eta_\xi+\frac{||\boldsymbol{w}-\boldsymbol{X}\boldsymbol{E}||_2^2}{2}\right\rangle_{\tilde{q}(/v_\xi)},
\end{aligned}
$$

and

$$
\begin{aligned}
\langle ||\boldsymbol{w}-\boldsymbol{X}\boldsymbol{E}||_2^2\rangle_{\tilde{q}(/v_\xi)} &= \sum_{\boldsymbol{i}}\Bigg(|\widetilde{m}_w(\boldsymbol{i})|^2+\widetilde{v}_w(\boldsymbol{i})+\Big(|\widetilde{m}_\chi(\boldsymbol{i})|^2+\widetilde{v}_w(\boldsymbol{i})\Big)\overline{E^2(\boldsymbol{i})} \\
&- 2\Re e\Big(\widetilde{m}_\chi(\boldsymbol{i})\overline{w^*(\boldsymbol{i})E(\boldsymbol{i})}\Big)\Bigg),
\end{aligned}
$$

By identification we find:

$$
\begin{cases}
\widetilde{\phi}_\xi^n &= \left(1-\alpha_{v_\xi}\right)\widetilde{\phi}_\xi+\alpha_{v_\xi}\left(\phi_\xi+\frac{N_PN}{2}\right)\left(\alpha_{v_\xi}\;\boldsymbol{d}_\chi+\left(1-\alpha_{v_\xi}\right)\widetilde{\boldsymbol{m}}_\chi\right) \\
\widetilde{\eta}_\xi^n &= \frac{1-\alpha_{v_\xi}}{\widetilde{\eta}_\xi}+\alpha_{v_\xi}\Bigg(\eta_\xi+\frac{1}{2}\Big(||\widetilde{\boldsymbol{m}}_w||_2^2+||\widetilde{\boldsymbol{V}}_w||_2^2+||\left(\widetilde{\boldsymbol{m}}_\chi^2+\widetilde{\boldsymbol{V}}_\chi\right)\overline{\boldsymbol{E}^2}||_1 \\
&- 2\Re e\left(\widetilde{\boldsymbol{m}}_\chi^H\overline{\boldsymbol{w}\circ\boldsymbol{E}^*}\right)\Big)\Bigg)
\end{cases} \tag{16}
$$

where $\boldsymbol{d}_\chi=\sum_k\overline{v_k^{-1}}\boldsymbol{\zeta}_k\circ\widetilde{\boldsymbol{m}}_\chi^\dagger+\overline{v_\xi^{-1}}\overline{\boldsymbol{w}\circ\boldsymbol{E}^*}$.

2.1.6 Variance of the classes v_k

$\forall\kappa\in\{1,...,K\}$ we have:

$$
\begin{aligned}
\log\left(\widetilde{q}^n(v_k)\right) &\propto (1-\alpha)\log\left(\widetilde{q}(v_k)\right) \;+\; \alpha_{v_k}\; \langle \log\left(p\left(\boldsymbol{\chi}, \boldsymbol{w}, \boldsymbol{z}, \boldsymbol{\psi}, \boldsymbol{y}\right)\right)\rangle_{\tilde{q}(/v_k)} \\
&\propto (1-\alpha_{v_k})\left(-\frac{v_k}{\tilde{\eta}_k} + (\tilde{\phi}_k - 1)\log(v_k)\right) + \alpha_{v_k}\left\langle -\frac{1}{2}\sum_{\boldsymbol{i}} \delta(z(\boldsymbol{i}) = \kappa)\right. \\
&\times \left.\left(\log(v_k) + \frac{|\chi(\boldsymbol{i}) - \boldsymbol{m}_k|^2}{v_k}\right) + \frac{\eta_0}{v_k} + (\phi_0 + 1)\log(v_k)\right\rangle_{\tilde{q}(/v_k)} \\
&\propto (1-\alpha_{v_k})\left[-\frac{v_k}{\tilde{\eta}_k} + (\tilde{\phi}_k - \log(v_k))\right] + \alpha_{v_k}\left[-\left(\phi_0 + 1 + \frac{\sum_{\boldsymbol{i}}\widetilde{\zeta}_k(\boldsymbol{i})}{2}\right)\right. \\
&- \left.\overline{v_k^{-1}}\left\langle \eta_0 + \frac{|\chi(\boldsymbol{i}) - \boldsymbol{m}_k|^2}{2}\right\rangle_{\tilde{q}(/v_k)}\right],
\end{aligned} \tag{17}
$$

$q(/v_k)) = q(\boldsymbol{\chi})q(\boldsymbol{w}))q(\boldsymbol{z})\prod_{l\neq v_k} q(\psi_l)$.

By identification we find:

$$
\left\{
\begin{array}{l}
\widetilde{\phi}_k^n = (1-\alpha_{v_k})\,\widetilde{\phi}_k + \alpha_{v_k}\left(\phi_0 + \frac{\sum_{\boldsymbol{i}}\tilde{\zeta}_k(\boldsymbol{i})}{2}\right) \\
\widetilde{\eta}_k^n = \frac{1-\alpha_{v_k}}{\widetilde{\eta}_k} + \alpha_{v_k}\left(\eta_0 + \frac{1}{2}\sum_{\boldsymbol{i}} \widetilde{\zeta}_k(\boldsymbol{i})\left(|\widetilde{m}_\chi(\boldsymbol{i})|^2 + \widetilde{V}_\chi(\boldsymbol{i}) + m_k^2 + \tilde{\tau}_k^2\right)\right)
\end{array}
\right. \tag{18}
$$

2.1.7 Means of classes m_k

$$
\begin{aligned}
\log\left(\widetilde{q}^n(m_k)\right) &\propto (1-\alpha_{m_k})\log\left(\widetilde{q}(m_k)\right) \;+\; \alpha_{m_k}\; \langle \log\left(p\left(\boldsymbol{\chi}, \boldsymbol{w}, \boldsymbol{z}, \boldsymbol{\psi}, \boldsymbol{y}\right)\right)\rangle_{\tilde{q}(/m_k)} \\
&\propto (1-\alpha_{m_k})\left[\widetilde{\tau}_k m_k^2 - 2\widetilde{\mu}_k m_k\right] + \alpha_{m_k}\left(-\left\langle\frac{1}{2}\sum_{\boldsymbol{i}} \delta\left(z(\boldsymbol{i}) = \kappa\right)\right.\right. \\
&\times \left.\left.\left(\frac{|\chi(\boldsymbol{i}) - m_\chi(\boldsymbol{i})|^2}{v_k}\right) + \frac{|m_k - \mu_0|}{\tau_0}\right\rangle_{\tilde{q}(/m_k)}\right),
\end{aligned} \tag{19}
$$

with $q(/m_k)) = q(\boldsymbol{\chi})q(\boldsymbol{w}))q(\boldsymbol{z})\prod_{l\neq m_k} q(\psi_l)$

By identification, we obtain:

$$
\left\{
\begin{array}{l}
\widetilde{\tau}_k^n = \left[(1-\alpha_{m_k})\,\widetilde{\tau}_k + \alpha_{m_k}\left(\tau_0^{-1} + \overline{v_k^{-1}}\sum_{\boldsymbol{i}}\widetilde{\zeta}_k(\boldsymbol{i})\right)\right]^{-1} \\
\widetilde{\mu}_k^n = \widetilde{\tau}_k^n\left[(1-\alpha_{m_k})\,\widetilde{\mu}_k + \alpha_{m_k}\left(\widetilde{\tau}_k\left(\frac{\mu_0}{\tau_0} + \overline{v_k^{-1}}\sum_{\boldsymbol{i}}\widetilde{\zeta}_k(\boldsymbol{i})\widetilde{m}_\chi(\boldsymbol{i})\right)\right)\right]
\end{array}
\right. \tag{20}
$$

3 The free negative energy

The evaluation of the free negative energy during the iteration process allows to have an indicator on the convergence of the algorithm. Indeed, its value at the convergence allows for an estimate of the evidence that the model is useful for the choice of the latter. In the nonlinear case, the free negative energy is given by:

$$
\mathcal{F}(\boldsymbol{q}) = \langle p(\boldsymbol{y}, \boldsymbol{\chi}, \boldsymbol{w}, \boldsymbol{z}, \boldsymbol{\psi}|\mathcal{M})\rangle_{\boldsymbol{q}} + \mathcal{H}(q(\boldsymbol{\chi}, \boldsymbol{w}, \boldsymbol{z}, \boldsymbol{\psi})) \tag{21}
$$

where $\mathcal{H}$ is the entropy of $\boldsymbol{q}$ and is given by:

$$
\begin{aligned}
\mathcal{H}(q(\boldsymbol{\chi},\boldsymbol{w},\boldsymbol{z},\boldsymbol{\psi})) &= \sum_{N_v}\sum_{N_f}\sum_{N_P}\sum_{\boldsymbol{i}}\sum_{k}\log\left(\sqrt{2\pi e\widetilde{v}_w(\boldsymbol{i})}\right)+\sum_{\boldsymbol{i}}\sum_{k}\log\left(\sqrt{2\pi e\widetilde{v}_\chi(\boldsymbol{i})}\right)\\
&+ \sum_{\boldsymbol{i}}\sum_{k}\tilde{\zeta}_\kappa(\boldsymbol{i})\left(\lambda\sum_{\boldsymbol{j}\in\mathcal{V}(\boldsymbol{i})}\tilde{\zeta}_k(\boldsymbol{j})-\log\left(\sum_k\exp\left\{\lambda\sum_{\boldsymbol{j}\in\mathcal{V}(\boldsymbol{i})}\tilde{\zeta}_k(\boldsymbol{j})\right\}\right)\right)\\
&+ \sum_{\kappa}\log\left(\sqrt{2\pi e\widetilde{\tau}_\kappa}\right)+\sum_k\left(\widetilde{\eta}_k+\log\left(\widetilde{\phi}_k\Gamma(\widetilde{\eta}_k)\right)-(1+\widetilde{\eta}_k)\Psi(\widetilde{\eta}_k)\right)\\
&+ \widetilde{\eta}_\epsilon+\log\left(\widetilde{\phi}_\epsilon\Gamma(\widetilde{\eta}_\epsilon)\right)-(1+\widetilde{\eta}_\epsilon)\Psi(\widetilde{\eta}_\epsilon)+\widetilde{\eta}_\xi+\log\left(\widetilde{\phi}_\xi\Gamma(\widetilde{\eta}_\xi)\right)\\
&- (1+\widetilde{\eta}_\xi)\Psi(\widetilde{\eta}_\xi),
\end{aligned}
\tag{22}
$$

Hence:

$$
\begin{aligned}
\langle p(\boldsymbol{y},\boldsymbol{\chi},\boldsymbol{w},\boldsymbol{z},\boldsymbol{\psi})\rangle_{\boldsymbol{q}} &= -\frac{M+N(1+N_P)+K}{2}\log(2\pi)+(\phi_\epsilon-1)\log\left(\eta_\epsilon\right)-\Gamma(\phi_\epsilon)\\
&- \frac{M+2\phi_\epsilon+2}{2}\left(\Psi(\widetilde{\eta}_\epsilon)-\log(\widetilde{\phi}_\epsilon)\right)-\eta_\epsilon\overline{v_\epsilon^{-1}}+(\phi_\xi-1)\log\left(\eta_\xi\right)-\Gamma(\phi_\xi)\\
&- \frac{NN_vN_PN_f+2\phi_\xi+2}{2}\left(\Psi(\widetilde{\eta}_\xi)-\log(\widetilde{\phi}_\xi)\right)-\eta_\xi\overline{v_\xi^{-1}}+\log(\eta_0^{K(\phi_0-1)})\\
&- \sum_k\frac{\left(2\phi_0+1+\sum_{\boldsymbol{i}}\widetilde{\zeta}_k(\boldsymbol{i})\right)\left(\Psi(\widetilde{\eta}_k)-\log(\widetilde{\phi_k})\right)-2\eta_0\overline{v_k^{-1}}}{2}-K\,\Gamma(\phi_0)\\
&- \sum_k\frac{\widetilde{m}_k^2+\widetilde{\tau}_k+\mu_0^2-2\widetilde{m}_k\mu_0}{2\tau_0}+\lambda\sum_{\boldsymbol{i}}\sum_k\widetilde{\zeta}_k(\boldsymbol{i})\sum_{\boldsymbol{j}\in\mathcal{V}(\boldsymbol{i})}\widetilde{\zeta}_k(\boldsymbol{j})-\frac{\overline{v_\epsilon^{-1}}}{2}\\
&\times \left(||\boldsymbol{y}||_2^2+||\boldsymbol{G}^o\widetilde{\boldsymbol{m}}_w||_2^2-2\boldsymbol{y}^t\boldsymbol{G}^o\widetilde{\boldsymbol{m}}_w+||\widetilde{\boldsymbol{V}}_w\Gamma^o||_1\right)\\
&- \frac{\overline{v_\xi^{-1}}}{2}\left(||\widetilde{\boldsymbol{m}}_w||_2^2+||\widetilde{\boldsymbol{V}}_w||_1+\left|\left|\left(\widetilde{\boldsymbol{m}}_\chi+\widetilde{\boldsymbol{V}}_\chi\right)\overline{\boldsymbol{E}^2}\right|\right|_1-2\widetilde{\boldsymbol{m}}_\chi^H\overline{\boldsymbol{w}\circ\boldsymbol{E}^*}\right)\\
&- \sum_{\boldsymbol{i}}\sum_k\frac{\overline{v_k^{-1}}\widetilde{\zeta}_k(\boldsymbol{i})\left(|\widetilde{m}_\chi(\boldsymbol{i})|_2^2+\widetilde{v}_\chi(\boldsymbol{i})+\widetilde{m}_\chi^{\dagger 2}(\boldsymbol{i})-2\widetilde{m}_\chi(\boldsymbol{i})\widetilde{m}_\chi^\dagger(\boldsymbol{i})\right)}{2},
\end{aligned}
\tag{23}
$$

We may note that the computation of the free negative energy depends mainly in terms used to update the values of the parameters and its evaluation does not require a cost of additional computation.

4 Optimal step values computation

By examining equations (9), (11), (12), (14), (16), (18), (20), we remark that there is no dependence between elements of the same group for χ, z, v_ϵ, v_ξ, v_k and m_k, whereas the update of the mean value of the contrast source $\widetilde{\boldsymbol{m}}_w^n$ at the iteration n depends on the mean value $\widetilde{\boldsymbol{m}}_w$. Then the gradient steps α_ρ, $\rho=\chi$, z, v_ϵ, v_ξ, m_k, v_k can be set to 1 in order to accelerate the convergence and only the contrast source updating step α_w is computed. This step is computed in an optimal way in order to ensure a fast convergence of the shaping parameter towards their final values. This is done from the negative free-energy function by means of a Newton's method. The optimal step then reads:

$$
\alpha_w^{opt}=\Delta\mathcal{F}(\alpha_w)/\Delta^2\mathcal{F}(\alpha_w)\Big|_{\alpha_w=0},
\tag{24}
$$

with $\Delta\mathcal{F}=\partial\mathcal{F}/\partial\alpha_w$ et $\Delta^2\mathcal{F}=\partial^2\mathcal{F}/\partial^2\alpha_w$

Before applying the Newton method, we rewrite the update equations for $\boldsymbol{w}$ with more convenient notations:

$$\begin{cases} \widetilde{\boldsymbol{V}}_w^n = \left[(1-\alpha_w)\, \widetilde{\boldsymbol{V}}_w^{-1} + \alpha\ \boldsymbol{R}_w \right]^{-1} \\ \widetilde{\boldsymbol{m}}_w^n = \widetilde{\boldsymbol{m}}_w + \alpha_w\ \widetilde{\boldsymbol{V}}_w^n\ \boldsymbol{d}_w \end{cases} \tag{25}$$

where $\boldsymbol{R}_w = \mathbf{Diag}\left(\overline{v_\epsilon^{-1}}\Gamma^o + \overline{v_\xi^{-1}}\overline{\Gamma^{xc}}\right)$ and $\boldsymbol{d}_w = \left[\overline{v_\epsilon^{-1}}\boldsymbol{G}^{o^H}(\boldsymbol{y} - \boldsymbol{G}^o\widetilde{\boldsymbol{m}}_w) + \overline{v_\xi^{-1}}\Big(\widetilde{\boldsymbol{m}}_\chi \boldsymbol{E}^{inc} - \boldsymbol{G}^{c^H}\left(\widetilde{\boldsymbol{m}}_\chi^2 + \widetilde{\boldsymbol{V}}_\chi\right)\boldsymbol{E}^{inc}\right.$
$\left. - \widetilde{\boldsymbol{m}}_w + \widetilde{\boldsymbol{m}}_\chi \boldsymbol{G}^c\widetilde{\boldsymbol{m}}_w + \boldsymbol{G}^{c^H}\widetilde{\boldsymbol{m}}_\chi^*\widetilde{\boldsymbol{m}}_w - \boldsymbol{G}^{c^H}\left(\widetilde{\boldsymbol{m}}_\chi^2 + \widetilde{\boldsymbol{V}}_\chi\right)\boldsymbol{G}^c\widetilde{\boldsymbol{m}}_w\Big)\right].$

Then, we rewrite the negative free energy as a function of α_w in order to use it to obtain the optimal value. It reads:

$$\begin{aligned} \mathcal{F}(\alpha_w) = &-\frac{\overline{v_\epsilon^{-1}}}{2}\left(||\boldsymbol{y}||_2^2 + ||\boldsymbol{G}^o\widetilde{\boldsymbol{m}}_w||_2^2 - 2\boldsymbol{y}^t\boldsymbol{G}^o\widetilde{\boldsymbol{m}}_w + ||\widetilde{\boldsymbol{V}}_w\Gamma^o||_1\right) - \frac{\overline{v_\xi^{-1}}}{2}\left(||\widetilde{\boldsymbol{m}}_w||_2^2 + ||\widetilde{\boldsymbol{V}}_w||_1\right. \\ &\left. + \left|\left|\left(\widetilde{\boldsymbol{m}}_\chi + \widetilde{\boldsymbol{V}}_\chi\right)\overline{\boldsymbol{E}^2}\right|\right|_1 - 2\widetilde{\boldsymbol{m}}_\chi^H\overline{\boldsymbol{w}\circ\boldsymbol{E}^*}\right) + \sum_{\boldsymbol{i}}\log\left(\sqrt{2\pi e\widetilde{v}_w(\boldsymbol{i})}\right), \end{aligned} \tag{26}$$

Then, we calculate the first and the second derivatives that yield the following:

$$\frac{\partial\mathcal{F}(\alpha_w)}{\partial\alpha_w} = \widetilde{\boldsymbol{m}}_w'\ \boldsymbol{d}_w^t + \sum_i \frac{\widetilde{v}_w'(\boldsymbol{i})}{2\ \widetilde{v}_w(\boldsymbol{i})} - \widetilde{v}_w'(\boldsymbol{i})r_w(\boldsymbol{i}), \tag{27}$$

$$\frac{\partial^2\mathcal{F}(\alpha_w)}{\partial^2\alpha_w} = \widetilde{\boldsymbol{m}}_w''\ \boldsymbol{d}_w^t + \widetilde{\boldsymbol{m}}_w'\ \boldsymbol{d}_w^{t'} + \sum_i \frac{\widetilde{v}_w''(\boldsymbol{i})\widetilde{v}_w(\boldsymbol{i}) - (\widetilde{v}_w'(\boldsymbol{i}))^2}{2\left(\widetilde{v}_w(\boldsymbol{i})\right)^2} - \widetilde{v}_w''(\boldsymbol{i})r_w(\boldsymbol{i}). \tag{28}$$

where $v_w(\boldsymbol{i})$ and $r_w(\boldsymbol{i})$ stand respectively for diagonal elements $\boldsymbol{i}$ of matrix $\boldsymbol{V}_w$ and $\boldsymbol{R}_w$, and

$$\begin{cases} \widetilde{\boldsymbol{V}}_w'^n = \left(\widetilde{\boldsymbol{V}}_w^n\right)^2\left(\widetilde{\boldsymbol{V}}_w^{-1} - \boldsymbol{R}_w\right) \\ \widetilde{\boldsymbol{V}}_w''^n = \left(\widetilde{\boldsymbol{V}}_w^n\right)^3\left(\widetilde{\boldsymbol{V}}_w^{-1} - \boldsymbol{R}_w\right)^2 \\ \widetilde{\boldsymbol{m}}_w'^n = \widetilde{\boldsymbol{V}}_w^n\ \boldsymbol{d}_w + \alpha\ \widetilde{\boldsymbol{V}}_w'^n\boldsymbol{d}_w \\ \widetilde{\boldsymbol{m}}_w''^n = 2\ \widetilde{\boldsymbol{V}}_w'^n\boldsymbol{d}_w + \alpha\ \widetilde{\boldsymbol{V}}_w''^n\boldsymbol{d}_w \end{cases} \tag{29}$$

then the optimal value in this case becomes

$$\alpha_w^{opt} = \frac{\boldsymbol{d}_w^t\widetilde{\boldsymbol{m}}_{w0}' + \frac{1}{2}\sum_i s_{w_i}^2}{\widetilde{\boldsymbol{m}}_{w0}''\ \boldsymbol{d}_w^t + \widetilde{\boldsymbol{m}}_{w0}'\ \boldsymbol{d}_w^{t'} + \frac{1}{2}\sum_i s_{w_i}^3} \tag{30}$$

where

$$\begin{cases} \left.\frac{\partial\widetilde{\boldsymbol{m}}_w}{\partial\alpha_w}\right|_{\alpha_w=0} = \widetilde{\boldsymbol{m}}_{w0}' = \widetilde{\boldsymbol{V}}_{w0}\boldsymbol{d}_w \\ \left.\frac{\partial^2\widetilde{\boldsymbol{m}}_w}{\partial^2\alpha_w}\right|_{\alpha_w=0} = \widetilde{\boldsymbol{m}}_{w0}'' = 2\ \widetilde{\boldsymbol{V}}_{w0}'\boldsymbol{d}_w \\ \left.\frac{\partial\widetilde{\boldsymbol{V}}_w}{\partial\alpha_w}\right|_{\alpha_w=0} = \widetilde{\boldsymbol{V}}_{w0}' = -\widetilde{\boldsymbol{V}}_{w0}S_w \end{cases} \tag{31}$$

and

$$\begin{aligned} \boldsymbol{d}_w^{t'} &= \widetilde{\boldsymbol{m}}_w'^t\left(\overline{v_\epsilon^{-1}}\boldsymbol{G}^o\boldsymbol{G}^H - \overline{v_\xi^{-1}}\Big(1 - \widetilde{\boldsymbol{m}}_\chi\boldsymbol{G}^c - \boldsymbol{G}^{c^H}\widetilde{\boldsymbol{m}}_\chi^* - \boldsymbol{G}^{c^H}(\widetilde{\boldsymbol{m}}_\chi + \widetilde{\boldsymbol{V}}_\chi)\boldsymbol{G}^c\Big)\right)^t, \\ s_{w_i} &= r_{w_i}\widetilde{v}_{w0_i} - 1. \end{aligned} \tag{32}$$

References

[1] V. Smídl and A. Quinn, *The Variational Bayes Method in Signal Processing*, Springer Verlag, Berlin, 2006.

[2] H. Ayasso, B. Duchêne, and A. Mohammad-Djafari, "Optical diffraction tomography within a variational Bayesian framework," *Inverse Problems in Science and Engineering*, vol. 20, no. 1, pp. 59–73, 2012.

[3] L. Gharsalli, H. Ayasso, B. Duchêne, and A. Mohammad-Djafari, "Microwave tomography for breast cancer detection within a variational Bayesian approach," in *IEEE European Signal Processing Conference (EUSIPCO)*, Marrakech, Marocco, 2013.

[4] L. Gharsalli, H. Ayasso, B. Duchêne, and A. Mohammad-Djafari, "A gradient-like variational Bayesian approach: application to microwave imaging for breast tumor detection," submitted in *IEEE International Conference on Image Processing (ICIP)*, France, Paris, 2014.

[5] A. Fraysse and T. Rodet, "A measure-theoretic variational bayesian algorithm for large dimensional problems," Tech. Rep. hal-00702259, http://hal.archives-ouvertes.fr/docs/00/70/22/59/PDF/var_bayV8.pdf, 2012.

[6] W. C. Gibson, *The Method of Moments in Electromagnetics*, Chapman & Hall/CRC, Boca Raton, 2007.