Parameter Estimation of the Normal Ratio Distribution with Variational Inference

Nizar BOUHLEL[1], Félix MERCIER[2], Angelina EL GHAZIRI[1] and David ROUSSEAU[2]
[1] *Institut Agro, Univ Angers, INRAE, IRHS, SFR QuaSaV, 49000, Angers, France*
{nizar.bouhlel, angelina.elghaziri}@agrocampus-ouest.fr
[2] *Université d'Angers, LARIS, UMR INRAe IRHS, Angers, France*
felix.mercier@orange.fr, david.rousseau@univ-angers.fr

***Abstract*—In this paper, we propose a new method for parameter estimation of the probability density function of the photosystem II (PSII) index in chlorophyll fluorescence imaging. The PSII index is modeled as the ratio of two normal distributions. The proposed method is based on hierarchical Bayesian modeling, and the mean field variational Bayes is performed to approximate the hierarchical Bayesian inference.The approach is evaluated using data acquired on *Arabidopsis thaliana*. The optimal variational Bayes posterior distributions are computed and then used to estimate the parameters. The preliminary results on the parameter estimation are satisfactory and meet our expectations.**

***Index Terms*—Chlorophyll fluorescence imaging, maximum quantum yield of photosystem II (PSII), hierarchical Bayesian model, parameter estimation, variational Bayes,**

I. Introduction

In the past two decades, chlorophyll fluorescence (ChlFl) imaging has been widely used for plant phenotyping [1], [2], [3], [4], [5]. It helps quantifying plant resistance, monitoring plant photosynthesis, and tracking disease development and leaf growth [6]. For instance, in ChlFl imaging technique, several fluorescence parameters are computed to track the photosynthesis of the plant [7], [8]. The maximum quantum yield of photosystem II (PSII) photochemistry $((F_m - F_0)/F_m)$ [9] is among the most used ChlFl parameters. PSII is an important indicator of plant stress [4]. This ratio depends on two measured fluorescence parameters, F_m maximum fluorescence yield and F_0 minimum fluorescence yield. The PSII serves as a biomarker to assess the normal or abnormal photosynthetic activity of plant tissue, thus making it possible to discriminate between diseased and healthy leaves.

In most of the literature, Gaussianity is assumed for the PSII. But, very recently the non-Gaussianity of the PSII has been highlighted empirically [10]. Moreover, the authors in [11] investigate the non-Gaussianity of the PSII by modeling the latter as the ratio of two normally distributed variables. In addition, the authors propose an Expectation-Maximization (EM) algorithm to estimate the parameters of the probability density function (pdf) of the ratio. In this work, we propose another parameter estimation method for the ratio by adopting a hierarchical Bayesian modeling. The variational inference is performed here in this study to approximate the hierarchical Bayesian inference and to provide an efficient estimation.

The paper is organized as follows. We start by introducing in section II the statistical modeling of the PSII as a ratio of two normal distributions. In section III, we estimate the ratio parameters using the mean field variational Bayes. We apply the proposed method to *arabidopsis thaliana* ChlFl imaging data, and we discuss the obtained results in section IV. Finally, some concluding remarks close up this paper.

II. Statistical modeling of photosystem II index

In the ChlFl imaging technique, the raw images F_m and F_0 are not directly used [7], [8]. Instead, they are combined to produce some indices which serve as a biomarker. We focus on the most common of these indices, known as the maximum quantum yield of PSII [9]:

$$\frac{F_v}{F_m} = 1 - \frac{F_0}{F_m}. \tag{1}$$

where F_v is the difference between F_m and the minimum fluorescence F_0. The PSII ratio is an indicator of plant stress and is among the most used ChlFl parameters.

In [11], it was shown that both F_m and F_0 are modeled as two independent Gaussian distributed variables. Consequently, the ratio distribution of F_v/F_m can be modeled in the following way. Let us consider the variables X and Y as F_0 and F_m respectively, where X and Y are two independent normally distributed variables given by

$$X \sim \mathcal{N}(\mu_x, \sigma_x^2) \quad \text{and} \quad Y \sim \mathcal{N}(\mu_y, \sigma_y^2). \tag{2}$$

The pdf of the variable $Z = X/Y$ is defined by [11]

$$f_Z(z) = \frac{\rho}{\pi(1+\rho^2 z^2)} e^{-\frac{1+\beta^2\rho^2}{2\delta_y^2}} \, {}_1F_1\left(1, \frac{1}{2}; \frac{1}{2\delta_y^2} \frac{(1+\beta\rho^2 z)^2}{1+\rho^2 z^2}\right) \tag{3}$$

where $\rho = \frac{\sigma_y}{\sigma_x}$, $\beta = \frac{\mu_x}{\mu_y}$, $\delta_y = \frac{\sigma_y}{\mu_y}$ and ${}_1F_1(.)$ is the confluent hypergeometric function, also known as Kummer's function defined as follows [12]:

$${}_1F_1(a; c; z) = \sum_{n=0}^{+\infty} \frac{(a)_n}{(c)_n} \frac{z^n}{n!}, \quad (a)_n\text{: the Pochhammer symbol}$$

As a consequence, the pdf of F_v/F_m is given by $f_Z(1-z)$.

III. Parameter Estimation of the Ratio of two Normal Distributions

We seek to estimate the set of parameters $\theta = \{\rho, \beta, \delta_y\}$ from observations $z = (z_1, ..., z_n)$, where n is the sample size, or otherwise, estimate $\theta = \{\mu_x, \sigma_x^2, \mu_y, \sigma_y^2\}$ since the former can be deduced from the latter. In the absence of an explicit solution of the maximum likelihood (3), as it is described in [11], the Expectation-Maximization (EM) algorithm is used to find an estimation $\hat{\theta}$ given a current estimate θ' of θ. In this study, we present another method to infer the parameters: the hierarchical Bayesian method performed by the variational Bayes (VB) method [13]. A hierarchical Bayesian model makes it possible to integrate all the existing knowledge on the phenomenon to be studied to compute expectations with respect to the posterior distribution. But often these multilevel models give rise to analytically intractable distributions. Consequently, sampling methods such as Markov Chain Monte Carlo (MCMC) are used. Some attractive tools for constructing an MCMC algorithm are the Metropolis-Hasting algorithm or the Gibbs sampler. More precisely, the posterior mean of θ is computed with respect to the posterior distribution $f(\theta|z)$ by sampling from it using a Gibbs sampler.

The variational inference is an optimization-based technique for approximate hierarchical Bayesian inference [14], and provides a computationally efficient alternative to the MCMC sampling methods when θ is high dimensional or fast computation is of primary interest. The VB approximates the posterior distribution by a distribution with density $q(\theta)$ belonging to some tractable family of distributions. The best VB approximation q^* is found by maximizing the variational lower bound $\mathcal{L}(q(\theta))$ on $\ln f(z)$ with respect to the posterior distribution so as to obtain the following solution:

$$q^* = \arg\max_{q(\theta)} \mathcal{L}(q(\theta)) = \arg\max_{q(\theta)} \int q(\theta) \ln \frac{f(\theta, z)}{q(\theta)} \mathrm{d}\theta \quad (4)$$

Maximizing the lower bound is equivalent to minimizing the Kullback-Leibler (KL) divergence from $q(\theta)$ to $f(\theta|z)$. The Mean Field VB (MFVB) is one of the classes of VB algorithms [13]. It can easily perform approximate Bayesian inference by assuming some form of factorization for q. Indeed, thanks to the mean field theory, the approximate posterior distribution is factorized into disjoint groups $q(\theta) = \prod_{i=1}^m q(\theta_i)$, thus breaking the dependencies among the random variables. With this approximation, the solution to the maximization problem in equation (4) can be derived as follows

$$\ln q^*(\theta_i) \propto E_{q(-\theta_i)}\{\ln f(\theta, z)\} \quad i = 1..m \quad (5)$$

where $E_{q(-\theta_i)}\{.\}$ denotes the expectation w.r.t. $q(\theta_1), ..., q(\theta_{i-1}), q(\theta_{i+1}), ..., q(\theta_m)$ defined by $E_{q(-\theta_i)}\{\ln f(\theta, z)\} = \int \ln f(\theta, z) \prod_{j\neq i} q(\theta_j) \mathrm{d}\theta_j$. In the following is the description of the approach.

A. Parameter Estimation with Mean Field Variational Bayes

We propose the hierarchical Bayesian model:

$$z_i|y_i, \mu_x, \sigma_x^2 \propto |y_i| f_X(z_i y_i) \quad (6)$$

$$y_i|\mu_y, \sigma_y^2 \sim \mathcal{N}(\mu_y, \sigma_y^2) \quad (7)$$

$$\mu_x \sim \mathcal{N}(\mu_{x0}, \sigma_{x0}^2) \quad (8)$$

$$\sigma_x^2 \sim \text{Inv-Gamma}(\alpha_{x0}, \beta_{x0}) \quad (9)$$

$$\mu_y \sim \mathcal{N}(\mu_{y0}, \sigma_{y0}^2) \quad (10)$$

$$\sigma_y^2 \sim \text{Inv-Gamma}(\alpha_{y0}, \beta_{y0}). \quad (11)$$

The conditional pdf of the observed data Z_i is derived based on the ratio representation in the ratio model $Z = X/Y$. Specifically, with the knowledge of the normally distributed X_i component and by performing the transformation from X_i to Z_i, the conditional pdf of the observed Z_i given the variable Y_i is obtained as $f(z_i|y_i, \mu_x, \sigma_x^2)$. The inverse Gamma distribution is selected as the pdf of the variables σ_x^2 and σ_y^2 since in literature, it has been shown that this distribution is effective in describing the variance [15]. The conjugate prior for the likelihood function $f(y_i|\mu_y, \sigma_y^2)$ is chosen to be a normal distribution with a fixed mean and variance, $(\mu_{y0}, \sigma_{y0}^2)$. The knowledge of the conjugate prior significantly simplifies the posterior distribution derivation. The joint pdf of all variables can be written when $\Lambda = (y, \mu_x, \sigma_x^2, \mu_y, \sigma_y^2)$ as

$$f(z, \Lambda) = f(z|y, \mu_x, \sigma_x^2) f(y|\mu_y, \sigma_y^2) f(\mu_x) f(\sigma_x^2) f(\mu_y) f(\sigma_y^2) \quad (12)$$

We derive the MFVB procedure for approximating the posterior distribution $f(y, \mu_x, \sigma_x^2, \mu_y, \sigma_y^2|z)$ by imposing the product of independant factors

$$q(y, \mu_x, \sigma_x^2, \mu_y, \sigma_y^2) = q(y) q(\mu_x) q(\sigma_x^2) q(\mu_y) q(\sigma_y^2). \quad (13)$$

B. Derivation of $q(y)$

The optimal VB posterior for y is given by

$$\ln q^*(y) = E_{q(\mu_x, \sigma_x^2, \mu_y, \sigma_y^2)}\{\ln f(z|y, \mu_x, \sigma_x^2) + \ln f(y|\mu_y, \sigma_y^2)\} + C$$
$$= \sum_{i=1}^n \ln |y_i| - \sum_{i=1}^n A_i y_i^2 + \sum_{i=1}^n B_i y_i + C \quad (14)$$

where A_i and B_i are defined as

$$A_i = \frac{1}{2}\left[E_{q(\sigma_x^2)}\{\frac{1}{\sigma_x^2}\} z_i^2 + E_{q(\sigma_y^2)}\{\frac{1}{\sigma_y^2}\}\right] \quad (15)$$

$$B_i = E_{q(\mu_x)}\{\mu_x\} E_{q(\sigma_x^2)}\{\frac{1}{\sigma_x^2}\} z_i + E_{q(\mu_y)}\{\mu_y\} E_{q(\sigma_y^2)}\{\frac{1}{\sigma_y^2}\}, \quad (16)$$

and C denoted the constants independent of y_i as they are unnecessary for identifying the optimal variational distribution $q^*(y)$ which is given as follows:

$$q^*(y) \propto \prod_{i=1}^n A_i \frac{|y_i| e^{-A_i y_i^2 + B_i y_i}}{{}_1F_1(1, \frac{1}{2}, \frac{B_i^2}{4A_i})}. \quad (17)$$

It is clear that $q^*(y)$ is a product of $q^*(y_i)$ with the moment of order 1 and 2 given as follows

$$\begin{cases} E_{q(y)}\{y_i\} = \frac{B_i}{A_i} \frac{{}_1F_1(2, \frac{3}{2}, \frac{B_i^2}{4A_i})}{{}_1F_1(1, \frac{1}{2}, \frac{B_i^2}{4A_i})} \\ E_{q(y)}\{y_i^2\} = \frac{1}{A_i} \frac{{}_1F_1(2, \frac{1}{2}, \frac{B_i^2}{4A_i})}{{}_1F_1(1, \frac{1}{2}, \frac{B_i^2}{4A_i})}. \end{cases} \quad (18)$$

Readers are referred to [11] for more details about the computation of the moment of order 1 and 2.

C. Derivation of $q(\mu_x)$ and $q(\sigma_x^2)$

The optimal VB posterior for μ_x is given by

$$\ln q^*(\mu_x) = E_{q(y,\sigma_x^2,\mu_y,\sigma_y^2)}\{\ln f(z|y,\mu_x,\sigma_x^2) + \ln f(\mu_x)\} + C$$
$$= -\frac{1}{2}A\left(\mu_x - \frac{B}{A}\right)^2 + C \quad (19)$$

with

$$A = nE_{q(\sigma_x^2)}\{\frac{1}{\sigma_x^2}\} + \frac{1}{\sigma_{x0}}, \quad (20)$$
$$B = \sum_{i=1}^{n} z_i E_{q(\sigma_x^2)}\{\frac{1}{\sigma_x^2}\}E_{q(y)}\{y_i\} + \frac{\mu_{x0}}{\sigma_{x0}^2}. \quad (21)$$

Here, C represents the constants independent of μ_x. They are unnecessary for identifying the optimal variational distribution $q^*(\mu_x)$. It follows that $q^*(\mu_x)$ is $\mathcal{N}(\mu_{xn}, \sigma_{xn}^2)$ distribution with parameters given as follows:

$$\sigma_{xn}^2 = 1/A = \left[nE_{q(\sigma_x^2)}\{\frac{1}{\sigma_x^2}\} + \frac{1}{\sigma_{x0}}\right]^{-1}, \quad (22)$$
$$\mu_{xn} = B/A = \sigma_{xn}^2\left[E_{q(\sigma_x^2)}\{\frac{1}{\sigma_x^2}\}\sum_{i=1}^{n} z_i E_{q(y)}\{y_i\} + \frac{\mu_{x0}}{\sigma_{x0}^2}\right]. \quad (23)$$

The optimal VB posterior for σ_x^2 is given by

$$\ln q^*(\sigma_x^2) = E_{q(y,\mu_x,\mu_y,\sigma_y^2)}\{\ln f(z|y,\mu_x,\sigma_x^2) + \ln f(\sigma_x^2)\} + C$$
$$= -(1+\alpha_{x0}+\frac{n}{2})\ln\sigma_x^2 - \frac{1}{\sigma_x^2}\times$$
$$\left[\beta_{x0} + \frac{1}{2}\sum_{i=1}^{n} E_{q(y,\mu_x)}\{(z_i y_i - \mu_x)^2\}\right]. \quad (24)$$

It follows that $q^*(\sigma_x^2)$ is Inv-Gam$(\alpha_{xn}, \beta_{xn})$ with parameters defined by

$$\alpha_{xn} = \alpha_{x0} + \frac{n}{2}, \quad (25)$$
$$\beta_{xn} = \beta_{x0} + \frac{1}{2}\sum_{i=1}^{n} z_i^2 E_{q(y)}\{y_i^2\} + \frac{n}{2}E_{q(\mu_x)}\{\mu_x^2\}$$
$$- E_{q(\mu_x)}\{\mu_x\}\sum_{i=1}^{n} z_i E_{q(y)}\{y_i\}. \quad (26)$$

After the identification of the optimal variational distributions $q^*(\mu_x)$ and $q^*(\sigma_x^2)$, we are able to compute the expectations w.r.t. $q^*(\mu_x)$ and $q^*(\sigma_x^2)$ as follows: $E_{q(\sigma_x^2)}\{\frac{1}{\sigma_x^2}\} = \frac{\alpha_{xn}}{\beta_{xn}}$, $E_{q(\mu_x)}\{\mu_x\} = \mu_{xn}$ and $E_{q(\mu_x)}\{\mu_x^2\} = \sigma_{xn}^2 + \mu_{xn}^2$.

D. Derivation of $q(\mu_y)$ and $q(\sigma_y^2)$

The optimal VB posterior for μ_y is given by:

$$\ln q^*(\mu_y) = E_{q(y,\mu_x,\sigma_x^2,\sigma_y^2)}\{\ln f(y|\mu_y,\sigma_y^2) + \ln f(\mu_y^2)\} + C$$
$$= -\frac{1}{2}A(\mu_y - \frac{B}{A})^2 + C \quad (27)$$

where

$$A = nE_{q(\sigma_y^2)}\{\frac{1}{\sigma_y^2}\} + \frac{1}{\sigma_{y0}^2}, \quad (28)$$
$$B = E_{q(\sigma_y^2)}\{\frac{1}{\sigma_y^2}\}\sum_{i=1}^{n} E_{q(y)}\{y_i\} + \frac{\mu_{y0}}{\sigma_{y0}^2}. \quad (29)$$

It follows that $q^*(\mu_y^2)$ is normal distribution $\mathcal{N}(\mu_{yn}, \sigma_{yn}^2)$ with parameters

$$\sigma_{yn}^2 = 1/A = \left[nE_{q(\sigma_y^2)}\{\frac{1}{\sigma_y^2}\} + \frac{1}{\sigma_{y0}}\right]^{-1}, \quad (30)$$
$$\mu_{yn} = B/A = \sigma_{yn}^2\left[E_{q(\sigma_x^2)}\{\frac{1}{\sigma_x^2}\}\sum_{i=1}^{n} E_{q(y)}\{y_i\} + \frac{\mu_{y0}}{\sigma_{y0}^2}\right]. \quad (31)$$

The optimal VB posterior for σ_y^2 is given by:

$$\ln q^*(\sigma_y^2) = E_{q(y,\mu_x,\sigma_x^2,\mu_y)}\{\ln f(y|\mu_y,\sigma_y^2) + \ln f(\sigma_y^2)\} + C$$
$$= -(1+\alpha_{y0}+\frac{n}{2})\ln\sigma_y^2 - \left[\beta_{y0} + \frac{1}{2}\sum_{i=1}^{n} E_{q(y,\mu_y)}\{(y_i-\mu_y)^2\}\right]\frac{1}{\sigma_y^2}.$$

It follows that $q^*(\sigma_y^2)$ is Inv-Gam$(\alpha_{yn}, \beta_{yn})$ with parameters defined by

$$\alpha_{yn} = \alpha_{y0} + \frac{n}{2} \quad (32)$$
$$\beta_{yn} = \beta_{y0} + \frac{1}{2}\sum_{i=1}^{n} E_{q(y)}\{y_i^2\} + \frac{n}{2}E_{q(\mu_y)}\{\mu_y^2\}$$
$$- E_{q(\mu_y)}\{\mu_y\}\sum_{i=1}^{n} E_{q(y)}\{y_i\}. \quad (33)$$

The determination of the optimal variational distributions $q^*(\mu_y)$ and $q^*(\sigma_y^2)$ allow us to compute the expectations w.r.t. $q^*(\mu_y)$ and $q^*(\sigma_y^2)$: $E_{q(\sigma_y^2)}\{\frac{1}{\sigma_y^2}\} = \frac{\alpha_{yn}}{\beta_{yn}}$, $E_{q(\mu_y)}\{\mu_y\} = \mu_{yn}$ and $E_{q(\mu_y)}\{\mu_y^2\} = \sigma_{yn}^2 + \mu_{yn}^2$.

The optimal variational distributions are then summarized as follows:

$$q^*(\mu_x) \sim \mathcal{N}\left(\frac{\frac{\mu_{x0}}{\sigma_{x0}^2} + \frac{\alpha_{xn}}{\beta_{xn}}\sum_{i=1}^{n} z_i E_{q(y)}(y_i)}{\frac{1}{\sigma_{x0}} + n\frac{\alpha_{xn}}{\beta_{xn}}}, \left[n\frac{\alpha_{xn}}{\beta_{xn}} + \frac{1}{\sigma_{x0}}\right]^{-1}\right) \quad (34)$$

$$q^*(\sigma_x^2) \sim \text{Inv-Gam}\Bigg(\alpha_{x0} + \frac{n}{2}, \beta_{x0} + \frac{1}{2}\sum_{i=1}^{n} z_i^2 E_{q(y)}\{y_i^2\}$$
$$+ \frac{n}{2}(\sigma_{xn}^2 + \mu_{xn}^2) - \mu_{xn}\sum_{i=1}^{n} z_i E_{q(y)}\{y_i\}\Bigg) \quad (35)$$

$$q^*(\mu_y) \sim \mathcal{N}\left(\frac{\frac{\mu_{y0}}{\sigma_{y0}^2} + \frac{\alpha_{yn}}{\beta_{yn}}\sum_{i=1}^{n} E_{q(y)}(y_i)}{\frac{1}{\sigma_{y0}} + n\frac{\alpha_{yn}}{\beta_{yn}}}, \left[n\frac{\alpha_{yn}}{\beta_{yn}} + \frac{1}{\sigma_{y0}}\right]^{-1}\right) \quad (36)$$

$$q^*(\sigma_y^2) \sim \text{Inv-Gam}\Bigg(\alpha_{y0} + \frac{n}{2}, \beta_{y0} + \frac{1}{2}\sum_{i=1}^{n} E_{q(y)}\{y_i^2\}$$
$$+ \frac{n}{2}(\sigma_{yn}^2 + \mu_{yn}^2) - \mu_{yn}\sum_{i=1}^{n} E_{q(y)}\{y_i\}\Bigg) \quad (37)$$

$$q^*(y_i) = A_i\frac{|y_i|e^{-A_i y_i^2 + B_i y_i}}{{}_1F_1(1, \frac{1}{2}, \frac{B_i^2}{4A_i})}, \text{ where } A_i = \frac{1}{2}\left[\frac{\alpha_{xn}}{\beta_{xn}}z_i^2 + \frac{\alpha_{yn}}{\beta_{yn}}\right]$$
$$\text{and } B_i = \mu_{xn}\frac{\alpha_{xn}}{\beta_{xn}}z_i + \mu_{yn}\frac{\alpha_{yn}}{\beta_{yn}}. \quad (38)$$

The parameters $(\mu_{xn}, \mu_{yn}, \sigma_{xn}^2, \sigma_{yn}^2, \alpha_{xn}, \beta_{xn}, \alpha_{yn}, \beta_{yn})$ are determined from Algorithm 1 with the following updating procedure:

Algorithm 1

Initialize: $(\mu_{xn}, \mu_{yn}) \in \mathbb{R}$, $(\sigma_{xn}^2, \sigma_{yn}^2) > 0$, $(\alpha_{xn}, \beta_{xn}) > 0$ and $(\alpha_{yn}, \beta_{yn}) > 0$

Update the following

For $i = 1..n$ (sample size)

$A_i = \frac{1}{2}\left[\frac{\alpha_{xn}}{\beta_{xn}} z_i^2 + \frac{\alpha_{yn}}{\beta_{yn}}\right]$

$B_i = \mu_{xn}\frac{\alpha_{xn}}{\beta_{xn}} z_i + \mu_{yn}\frac{\alpha_{yn}}{\beta_{yn}}$

$E_{q(y)}\{y_i\} = \frac{B_i}{A_i}\frac{{}_1F_1(2,\frac{3}{2},\frac{B_i^2}{4A_i})}{{}_1F_1(1,\frac{1}{2},\frac{B_i^2}{4A_i})}$

$E_{q(y)}\{y_i^2\} = \frac{1}{A_i}\frac{{}_1F_1(2,\frac{1}{2},\frac{B_i^2}{4A_i})}{{}_1F_1(1,\frac{1}{2},\frac{B_i^2}{4A_i})}$

End For

$\beta_{xn} \leftarrow \beta_{x0} + \frac{1}{2}\sum_{i=1}^n z_i^2 E_{q(y)}\{y_i^2\} + \frac{n}{2}(\sigma_{xn}^2 + \mu_{xn}^2)$ - $\mu_{xn}\sum_{i=1}^n z_i E_{q(y)}\{y_i\}$

$\sigma_{xn}^2 \leftarrow \left[n\frac{\alpha_{xn}}{\beta_{xn}} + \frac{1}{\sigma_{x0}}\right]^{-1}$

$\mu_{xn} \leftarrow \sigma_{xn}^2\left[\frac{\mu_{x0}}{\sigma_{x0}^2} + \frac{\alpha_{xn}}{\beta_{xn}}\sum_{i=1}^n z_i E_{q(y)}(y_i)\right]$

$\beta_{yn} \leftarrow \beta_{y0} + \frac{1}{2}\sum_{i=1}^n E_{q(y)}\{y_i^2\} + \frac{n}{2}(\sigma_{yn}^2 + \mu_{yn}^2) - \mu_{yn}\sum_{i=1}^n E_{q(y)}\{y_i\}$

$\sigma_{yn}^2 \leftarrow \left[n\frac{\alpha_{yn}}{\beta_{yn}} + \frac{1}{\sigma_{y0}}\right]^{-1}$

$\mu_{yn} \leftarrow \sigma_{yn}^2\left[\frac{\mu_{y0}}{\sigma_{y0}^2} + \frac{\alpha_{yn}}{\beta_{yn}}\sum_{i=1}^n E_{q(y)}(y_i)\right]$

Until convergence.

Return $(\mu_{xn}, \mu_{yn}, \sigma_{xn}^2, \sigma_{yn}^2, \alpha_{xn}, \beta_{xn}, \alpha_{yn}, \beta_{yn})$

We stop the iterative scheme when the change of the ℓ_2-norm of the vector $(\mu_{xn}, \sigma_{xn}^2, \mu_{yn}, \sigma_{yn}^2, \alpha_{xn}, \beta_{xn}, \alpha_{yn}, \beta_{yn})^T$ is smaller than $\epsilon = 10^{-3}$. We derive the posterior mean or the posterior mode based on the marginal posterior distribution $q^*(\mu_x)$, $q^*(\mu_y)$, $q^*(\sigma_x^2)$ and $q^*(\sigma_y^2)$. The mean or mode is the value at which the posterior pdf takes its maximum value. Based on VB distributions, we compute the posterior mean

$$E_{q^*(\mu_x)}(\mu_x) = \mu_{xn}; \text{ and } E_{q^*(\sigma_x^2)}(\sigma_x^2) = \frac{\beta_{xn}}{\alpha_{xn} - 1} \quad (39)$$

$$E_{q^*(\mu_y)}(\mu_y) = \mu_{yn}; \text{ and } E_{q^*(\sigma_y^2)}(\sigma_y^2) = \frac{\beta_{yn}}{\alpha_{yn} - 1}. \quad (40)$$

Finally we can deduce the values of the estimated parameters

$$\begin{cases} \hat{\beta} = \dfrac{E_{q^*(\mu_x)}(\mu_x)}{E_{q^*(\mu_y)}(\mu_y)} = \dfrac{\mu_{xn}}{\mu_{yn}} \\ \hat{\rho} = \sqrt{\dfrac{E_{q^*(\sigma_y^2)}(\sigma_y^2)}{E_{q^*(\sigma_x^2)}(\sigma_x^2)}} = \sqrt{\dfrac{\beta_{yn}(\alpha_{xn}-1)}{\beta_{xn}(\alpha_{yn}-1)}} \\ \hat{\delta}_y = \dfrac{\sqrt{E_{q^*(\sigma_y^2)}(\sigma_y^2)}}{E_{q^*(\mu_y)}(\mu_y)} = \dfrac{1}{\mu_{yn}}\sqrt{\dfrac{\beta_{yn}}{\alpha_{yn}-1}}. \end{cases} \quad (41)$$

IV. Experimental results and discussion

The ChlFl imaging on rosettes of *Arabidopsis thaliana* is considered here. The image of the PSII index is made available to conduct the parameter estimation approach. To build variational inference with the mean field approximation, we selected area located in the limb of the leaves as illustrated in Figure (1.a). The normalized histogram of the selected area is represented in Figure (1.b). The ChlFl experimenters have suggested the following parameter $\beta = 0.1429, \rho = 6.0000, \delta_y = 0.2143$ for purpose of comparison with our estimation method. We proceed by using the following values for the initialization of algorithm: $\alpha_{x0} = 1$; $\beta_{x0} = 500$; $\mu_{x0} = 10$; $\sigma_{x0}^2 = 1.5^2$; $\alpha_{y0} = 1$; $\beta_{y0} = 15000$; $\mu_{y0} = 42$; $\sigma_{y0}^2 = 4^2$. The number of samples are $n = 1000$. The MFVB algorithm is stopped after 460 iterations. Figure (2) show the evolution of the parameters $(\mu_{xn}, \mu_{yn}, \sigma_{xn}, \sigma_{yn}, \beta_{xn}, \beta_{yn})$ over iterations from application of Algorithm 1. It is clear to see the convergence of these parameters starting from iteration 200. Figure (3) show the approximate posterior density functions for the four model parameters $(\mu_x, \sigma_x^2, \mu_y, \sigma_y^2)$. The estimated parameters $(\hat{\beta}, \hat{\rho}, \hat{\delta}_y)$ over the iterations are shown in Figure (4). The horizontal lines with red color correspond to the true values of these parameters. The final values of these parameters are $\hat{\beta} = 0.1438, \hat{\rho} = 5.8782, \hat{\delta}_y = 0.2372$, which are very close to the true parameters. Figure (5) depicts the comparison between the normalized histogram of the index F_0/F_m and the estimated pdf $f_Z(z)$ computed based on the estimated parameters $(\hat{\beta}, \hat{\rho}, \hat{\delta}_y)$ over the selected area. The fit ability is evaluated qualitatively by using the Kolmogorov-Smirnov (KS) hypothesis test. Based on the p-value ($= 0.36$) we can confirm that the estimated pdf curve fit well the normalized histogram.

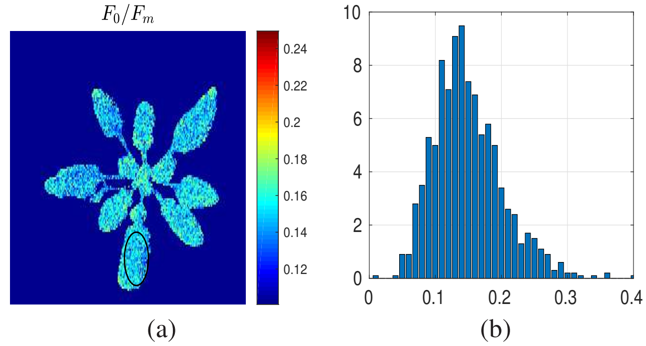

Fig. 1: Image of F_0/F_m (a). Plots of normalized histogram of the selected area in image F_0/F_m (b).

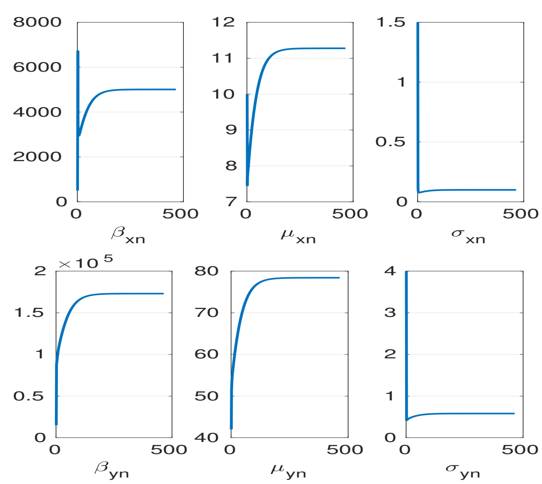

Fig. 2: Updates of $(\mu_{xn}, \mu_{yn}, \sigma_{xn}, \sigma_{yn}, \beta_{xn}, \beta_{yn})$ over iterations.

It is worth noticing that the algorithm is insensitive to the initialization of the parameters $(\beta_{xn}, \mu_{xn}, \sigma_{xn}^2, \beta_{yn}, \mu_{yn}, \sigma_{yn}^2)$ since the final result does not change. What really changes is the speed of convergence, more precisely, the number of iteration increases when the initial values are not chosen

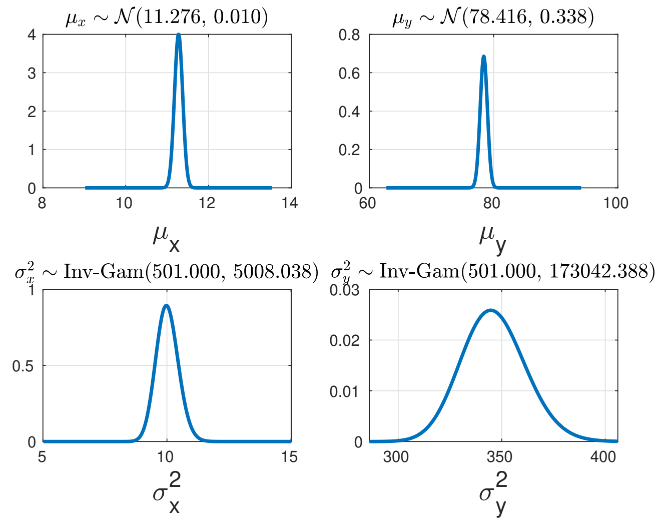

Fig. 3: Approximate marginal posterior density function for μ_x, σ_x^2, μ_y and σ_y^2 estimated by MFVB.

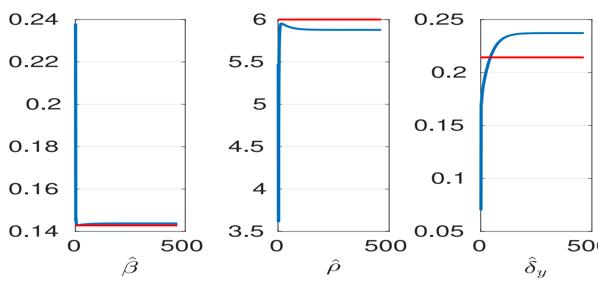

Fig. 4: Updates of $(\hat{\beta}, \hat{\rho}, \hat{\delta}_y)$ over iterations.

carefully. It should be noted that if we consider: $\mu'_x = a\mu_x$, $\sigma'_x = a\sigma_x$, $\mu'_y = a\mu_y$ and $\sigma'_y = a\sigma_y$ where $a > 0$ is a real constant then

$$\rho = \frac{\sigma'_y}{\sigma'_x} = \frac{\sigma_y}{\sigma_x}, \quad \beta = \frac{\mu'_x}{\mu'_y} = \frac{\mu_x}{\mu_y}, \quad \delta_y = \frac{\sigma'_y}{\mu'_y} = \frac{\sigma_y}{\mu_y} \tag{42}$$

This means that for different quadruplets $(\mu'_x, \sigma'_x, \mu'_y, \sigma'_y)$ and $(\mu_x, \sigma_x, \mu_y, \sigma_y)$ one can have the same parameters (β, ρ, δ_y), and consequently the same distribution $f_Z(z)$.

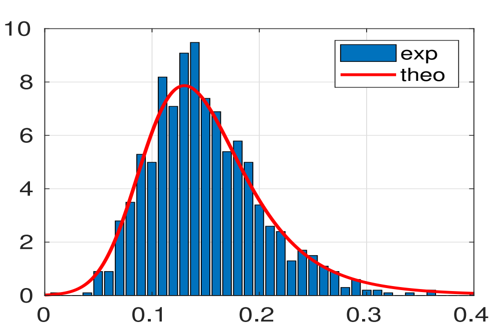

Fig. 5: Plot of PSII normalized histogram and estimated distribution given by (3) using the final estimated parameters $(\hat{\beta}, \hat{\rho}, \hat{\delta}_y)$.

V. Conclusion

A method for parameter estimation of the PSII index in chlorophyll fluorescence imaging has been proposed. The method is based on a hierarchical Bayesian modeling. The mean field approximation is performed to approximate the hierarchical Bayesian inference and to provide an efficient estimation. The approach has been evaluated on data acquired on rosettes of *Arabidopsis thaliana*. The estimation results are encouraging and can further be improved in several ways. The first way is by evaluating the performance for different sample sizes. A second way is by introducing some dependencies between μ_x and σ_x^2 in their models, the same process for μ_y and σ_y^2. We expect that these additional actions will enhance the estimation performance.

References

[1] E. Gorbe and A. Calatayud, "Applications of chlorophyll fluorescence imaging technique in horticultural research: A review," *Scientia Horticulturae*, vol. 138, pp. 24–35, 2012.

[2] H. M. Kalaji, G. Schansker, R. J. Ladle, V. Goltsev, K. Bosa, S. I. Allakhverdiev, M. Brestic, F. Bussotti, A. Calatayud, P. Dabrowski *et al.*, "Frequently asked questions about in vivo chlorophyll fluorescence: practical issues," *Photosynthesis research*, vol. 122, no. 2, pp. 121–158, 2014.

[3] H. M. Kalaji, G. Schansker, M. Brestic, F. Bussotti, A. Calatayud, L. Ferroni, V. Goltsev, L. Guidi, A. Jajoo, P. Li *et al.*, "Frequently asked questions about chlorophyll fluorescence, the sequel," *Photosynthesis Research*, vol. 132, no. 1, pp. 13–66, 2017.

[4] M. L. Pérez-Bueno, M. Pineda, and M. Barón, "Phenotyping plant responses to biotic stress by chlorophyll fluorescence imaging," *Frontiers in Plant Science*, vol. 10, 2019.

[5] R. Valcke, "Can chlorophyll fluorescence imaging make the invisible visible?" *Photosynthetica*, vol. 59, pp. 381–398, 2021.

[6] L. Li, Q. Zhang, and D. Huang, "A review of imaging techniques for plant phenotyping," *Sensors*, vol. 14, no. 11, pp. 20 078–20 111, 2014.

[7] S. Berger, Z. Benediktyová, K. Matouš, K. Bonfig, M. J. Mueller, L. Nedbal, and T. Roitsch, "Visualization of dynamics of plant-pathogen interaction by novel combination of chlorophyll fluorescence imaging and statistical analysis: differential effects of virulent and avirulent strains of p. syringae and of oxylipins on a. thaliana," *Journal of Experimental Botany*, vol. 58, no. 4, pp. 797–806, 2006.

[8] A. M. Sánchez-Moreiras, E. Graña, M. J. Reigosa, and F. Araniti, "Imaging of chlorophyll a fluorescence in natural compound-induced stress detection," *Frontiers in Plant Science*, vol. 11, 2020.

[9] B. Genty and S. Meyer, "Quantitative mapping of leaf photosynthesis using chlorophyll fluorescence imaging," vol. 22, no. 2, pp. 277–284, 1995.

[10] M. Pavicic, K. Overmyer, A. u. Rehman, P. Jones, D. Jacobson, and K. Himanen, "Image-based methods to score fungal pathogen symptom progression and severity in excised arabidopsis leaves," *Plants*, vol. 10, no. 1, 2021.

[11] A. El Ghaziri, N. Bouhlel, N. Sapoukhina, and D. Rousseau, "On the importance of non-gaussianity in chlorophyll fluorescence imaging," *Remote Sensing*, vol. 15, no. 2, 2023.

[12] I. Gradshteyn and I. Ryzhik, *Table of Integrals, Series, and Products, 7th Edition*. Academic Press is an imprint of Elsevier, 2007.

[13] M. P. Wand, J. T. Ormerod, S. A. Padoan, and R. Frühwirth, "Mean field variational bayes for elaborate distributions," *Bayesian Analysis*, vol. 6, pp. 847–900, 2011.

[14] J. T. Ormerod and M. P. Wand, "Explaining variational approximations," *The American Statistician*, vol. 64, no. 2, pp. 140–153, 2010.

[15] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, "Bayesian data analysis, third edition." Chapman and Hall/CRC, 2013.