Bayesian modeling of differentially expressed genes

March 25, 2008

1 Notation

$p \in \{1, \ldots, P\}$: study.

$g \in \{1, \ldots, G\}$: gene.

$s \in \{1, \ldots, S_p\}$: sample (array) in study p.

x_{gsp}: observed expression for gene g and sample s in study p.

$\psi_{sp} \in \{0, 1\}$ clinical variable for sample s in study p.

$\delta_{gp} \in \{0, 1\}$: indicator for differential expression for gene g and study p.

ν_{gp}: the true underlying mean expression (mean over $\psi_{sp} = 0$ and $\psi_{sp} = 1$).

Δ_{gp}: true differential expression (offsets) between the two clinical states.

$R = [R_{pq}]_{p,q=1}^{P}$: covariance matrix for $\boldsymbol{\Delta}_g = (\Delta_{g1}, \ldots, \Delta_{gP})^T$.

$\Sigma = [\Sigma_{pq}]_{p,q=1}^{P}$: covariance matrix for $\boldsymbol{\nu}_g = (\nu_{g1}, \ldots, \nu_{gP})^T$.

$\sigma^2_{g\psi p}$: variance of expression around the true mean.

2 Stochastic model

Let x_{gsp} denote observed expression value for gene $g \in \{1, \ldots, G\}$ and sample (array) $s \in \{1, \ldots, S_q\}$ in study $p \in \{1, \ldots, P\}$, where $P \geq 2$. We assume that some clinical variable (with two possible values) is available for each of the arrays in all studies. We denote this by $\psi_{sp} \in \{0, 1\}$ for sample (array) number s in study number p. We will assume that the studies have been somehow standardized so that the values in each study is centered around zero and are approximately Gaussian distributed.

Our model defined below is based on the following assumption, for some of the genes, the expression values x_{gsp} are differentially expressed (have different mean values) for arrays where $\psi_{sp} = 0$ and arrays where $\psi_{sp} = 1$. The differential expression can be for all studies or only some of them.

We assume the following hierarchical Bayesian model for the expression data. Conditionally on underlying unobserved parameters we assume the expression values to be independently Gaussian distributed,

$$x_{gsp}|\ldots \sim \mathrm{N}(\nu_{gp} + \delta_{gp}(2\psi_{sp} - 1)\Delta_{gp}, \sigma^2_{g\psi_{sp}p}), \tag{1}$$

where $\delta_{gp} \in \{0, 1\}$ indicates whether gene g is differentially expressed ($\delta_{gp} = 1$) or not ($\delta_{gp} = 0$) in study p. Thus, if $\delta_{gp} = 0$, the mean of x_{gsp} is ν_{gp}, whereas if $\delta_{gp} = 1$, the mean is $\nu_{gp} - \Delta_{gp}$ and $\nu_{gp} + \Delta_{gp}$ for samples with $\psi_{sp} = 0$ and $\psi_{sp} = 1$, respectively.

Given hyper-parameters, we assume $\boldsymbol{\nu}_g = (\nu_{g1}, \ldots, \nu_{gP})^T$ and $\boldsymbol{\Delta}_g = (\Delta_{g1}, \ldots, \Delta_{gP})^T$ to be multi-Gaussian distributed,

$$\boldsymbol{\nu}_g|\ldots \sim \mathrm{N}(0, \Sigma) \qquad \text{and} \qquad \boldsymbol{\Delta}_g|\ldots \sim \mathrm{N}(0, R), \tag{2}$$

where $\Sigma = [\Sigma_{pq}] \in \Re^{P\times P}$ and $R = [R_{pq}] \in \Re^{P\times P}$ with

$$\Sigma_{pq} = \gamma^2 \rho_{pq} \sqrt{\tau_p^2 \tau_q^2 \sigma_{gp}^{2a_p} \sigma_{gq}^{2a_q}} \tag{3}$$

and

$$R_{pq} = c^2 r_{pq} \sqrt{\tau_p^2 \tau_q^2 \sigma_{gp}^{2b_p} \sigma_{gq}^{2b_q}}. \tag{4}$$

The a_p and $b_p, p = 1, \ldots P$ are assumed apriori independent with discrete probabilities in 0 and 1 and a continuous density on $(0, 1)$. More precisely, we let

$$\mathrm{P}(a_p = 0) = p_a^0 \quad , \quad \mathrm{P}(a_p = 1) = p_a^1 \quad , \quad a_p|a_p \in (0, 1) \sim \mathrm{Beta}(\alpha_a, \beta_a), \tag{5}$$

and

$$\mathrm{P}(b_p = 0) = p_b^0 \quad , \quad \mathrm{P}(b_p = 1) = p_b^1 \quad , \quad b_p|b_p \in (0, 1) \sim \mathrm{Beta}(\alpha_b, \beta_b). \tag{6}$$

To c^2 we assign a uniform prior distribution on $[0, c^2_{\mathrm{Max}}]$, where c^2_{Max} is a parameter specified by the user, and for γ^2 we use an improper uniform distributions on $(0, \infty)$. We restrict $\tau_p^2 > 0, p = 1, \ldots, P$ and $\tau_1^2 \cdot \ldots \cdot \tau_P^2 = 1$ and assume an (improper) uniform distribution for $(\tau_1^2, \ldots, \tau_P^2)$ under this restriction. The $[\rho_{pq}] \in \Re^{P\times P}$ and $[r_{pq}] \in \Re^{P\times P}$ are restricted to be correlation matrices and we assign apriori independent Barnard et al (2000) distributions for them, with ν_ρ and ν_r degrees of freedom for $[\rho_{pq}]$ and $[r_{pq}]$, respectively.

For the indicators $\delta_{gp}, g = 1, \ldots, G, p = 1, \ldots, P$ we have implemented two prior models. In prior model A we assume, for each $p = 1, \ldots, P$, $\delta_{1p}, \ldots, \delta_{Gp}$ to be apriori independent, given a hyper-parameter ξ_p, with

$$\mathrm{P}(\delta_{gp} = 1|\xi_p) = \xi_p, \qquad \text{where} \qquad \xi_1, \ldots, \xi_P \sim \mathrm{Beta}(\alpha_\xi, \beta_\xi) \tag{7}$$

independently. In prior model B we set the restriction $\delta_{g1} = \ldots = \delta_{gP} = \delta_g$ for each $g = 1, \ldots, G$ and assume the G indicators $\delta_1, \ldots, \delta_G$ to be apriori independent, given a hyper-parameter ξ, with

$$\mathrm{P}(\delta_g = 1|\xi) = \xi, \qquad \text{where} \qquad \xi \sim \mathrm{Beta}(\alpha_\xi, \beta_\xi). \tag{8}$$

We assume σ^2_{gp} in (3) to be given from $\sigma^2_{g\psi p}$ in (1) via the following relations

$$\sigma^2_{g0p} = \sigma^2_{gp}\varphi_{gp} \qquad \text{and} \qquad \sigma^2_{g1p} = \frac{\sigma^2_{gp}}{\varphi_{gp}}. \tag{9}$$

For σ^2_{gp} we assume independent prior distributions, given hyper-parameters,

$$\sigma^2_{gp}|\ldots \sim \text{Gamma}\left(\frac{l_p^2}{t_p}, \frac{l_p}{t_p}\right), \tag{10}$$

where the mean l_p and variance t_p for $p = 1, \ldots, P$ are assigned independent (improper) uniform distributions on $(0, \infty)$. The φ_{gp} are assigned independent gamma priors, given hyper-parameters,

$$\varphi_{gp}|\ldots \sim \text{Gamma}\left(\frac{\lambda_p^2}{\theta_p}, \frac{\lambda_p}{\theta_p}\right), \tag{11}$$

where the mean λ_p and the variance θ_p for $p = 1, \ldots, P$ are assigned independent (improper) uniform distributions on $(0, \infty)$.

3 Metropolis–Hastings algorithm

To simulate from the posterior distribution resulting from the above specified Bayesian model we use a Metropolis–Hastings algorithm with the following moves.

1. The ν_{pg}'s are updated in a Gibbs step.

2. The Δ_{pg}'s are updated in a Gibbs step.

3. Each $a_p, p = 1, \ldots, P$ is updated in turn, where a potential new value for a_p, $\widetilde{a}_p$, is generated as follows. If $a_p = 0$, $\widetilde{a}_p \sim \text{Uniform}(0, \varepsilon)$, if $a_p = 1$, $\widetilde{a}_p \sim \text{Uniform}(1 - \varepsilon, 1)$, and if $a_p \in (0, 1)$ we use

$$\text{P}(\widetilde{a}_p = 0) = \max\{-(a_p - \varepsilon), 0\} \cdot \text{I}(p_a^0 \neq 0), \tag{12}$$

$$\text{P}(\widetilde{a}_p = 1) = \max\{a_p + \varepsilon - 1, 0\} \cdot \text{I}(p_a^1 \neq 0) \tag{13}$$

and

$$\widetilde{a}_p|\widetilde{a}_p \in (0, 1) \sim \text{Uniform}(\max\{a_p - \varepsilon, 0\}, \min\{a_p + \varepsilon, 1\}). \tag{14}$$

The ε is a tuning parameter that has to be specified by the user. The default value is 0.05.

4. Each $b_p, p = 1, \ldots, P$ is updated in turn, where the proposal distribution is of the same type as in 3. The default value for ε is again 0.05.

5. A block Gibbs update for c^2 and Δ_{gp} for (g, p)'s where $\delta_{gp} = 0$.

6. The γ^2 is updated in a Gibbs step.

7. A block update for the correlation matrix $[r_{pq}]$ and the variance c^2. First, potential new values for $[r_{pq}]$, $[\widetilde{r}_{pq}]$, is set by

$$\widetilde{r}_{pq} = (1-\varepsilon)r_{pq} + \varepsilon T_{pq}, \tag{15}$$

where $[T_{pq}]$ is a correlation matrix which with probability a half is generated from the prior for $[r_{pq}]$, or with probability a half set equal to unity on the diagonal and with all off diagonal elements set equal to the same value b. In the latter case, the value b is sampled from a uniform distribution on $(-1/(P-1), 1)$. Second, the potential new value for c^2 is sampled from the corresponding full conditional (given the potential new values $[\widetilde{r}_{pq}]$). The ε is a tuning parameter that has to be supplied by the user and its default value is 0.01.

8. A block update for the correlation matrix $[\rho_{pq}]$ and the variance γ^2. The potential new values are generated similar to what is done in 7. The default value for ε is again 0.01.

9. For each $g = 1, \ldots, G$ in turn, a block update for δ_{gp} and $\mathbf{\Delta}_g$ for a randomly chosen p. First, the potential new value for δ_{gp} is set equal to $\widetilde{\delta}_{gp} = 1 - \delta_{gp}$. Second, the potential new value for $\mathbf{\Delta}_g$ is sampled from the full conditional (given the potential new value $\widetilde{\delta}_{gp}$). No tuning parameter for this update.

10. The $\xi_1, \ldots, \xi_P$ (prior model A) or ξ (prior model B) are updated in a Gibbs step.

11. Each σ^2_{gp} is updated in turn, where the potential new value is given as $\widetilde{\sigma}^2_{gp} = \sigma^2_{gp} \cdot u$, where $u \sim \text{Uniform}(1/(1+\varepsilon), 1+\varepsilon)$. The ε is a tuning parameter that has to be specified by the user and its default value is 0.5.

12. Each $t_p, p = 1, \ldots, P$ is updated in turn. The potential new value is given as $\widetilde{t}_p = t_p \cdot u$, where $u \sim \text{Uniform}(1/(1+\varepsilon), 1+\varepsilon)$. The ε is a tuning parameter that has to be specified by the user and its default value is 0.1.

13. Each $l_p, p = 1, \ldots, P$ is updated in turn. The potential new value is given as $\widetilde{l}_p = l_p \cdot u$, where $u \sim \text{Uniform}(1/(1+\varepsilon), 1+\varepsilon)$. The ε is a tuning parameter that has to be specified by the user and its default value is 0.05.

14. Each φ_{gp} is updated in turn, where the proposal distribution is of the same type as in 11. The default value for ε is again 0.5.

15. Each $\theta_p, p = 1, \ldots, P$ is updated in turn, where the proposal distribution is of the same type as in 12. The default value for ε is again 0.1.

16. Each $\lambda_p, p = 1, \ldots, P$ is updated in turn, where the proposal distribution is of the same type as in 13. The default value for ε is again 0.05.

17. The $(\tau^2_1, \ldots, \tau^2_P)$ is updated by first uniformly at random drawing a pair $p, q \in \{1, \ldots, P\}$ so that $p \neq q$. Potential new values for τ^2_p and τ^2_q are generated by setting

$$\widetilde{\tau}^2_p = \tau^2_p \cdot u \quad \text{and} \quad \widetilde{\tau}^2_q = \tau^2_q / u, \tag{16}$$

where $u \sim \text{Uniform}(1/(1+\varepsilon), 1+\varepsilon)$. The ε is again a tuning parameter that has to be specified by the user and its default value is set to 0.02.

4 Input files

The computer program is started with the command *diffExpressed*. This command takes the following seven parameters.

1. Seed value. A positive integer used as seed value in the random number generator.
2. Number of iterations. A non-negative integer value specifying the number of iterations to run in the Metropolis–Hastings algorithm.
3. Prior model indicator, must be 0 or 1. If 0 is set, prior model A for δ_{gp} is used. Otherwise prior model B is used.
4. Data file. A filename. The file should specify the data set. See description of the format below.
5. Parameter file. A filename. The file may specify hyper-parameter values. See description of the format below.
6. Initial values file. A filename. The file may specify initial values for the parameter values. See description of the format below.
7. Update file. A filename. The file may specify the number of updates of each type in each iteration, and a value for the tuning parameter ϵ (for updates where this is relevant). See description of the format below.
8. Output file. A filename. The file may specify which simulated values to print to files. See description of the file format below.

For any of parameters 5 to 8 a '=' may be used instead of a filename. If so, default values will be used. For choice of default values, see description below.

In the following, we describe the format of each the five input files.

4.1 Data file

The first line of the data file should contain an integer, giving the number of studies, P. Each of the following P lines should give two file names. The first file should contain expression values, whereas the second file should contain corresponding clinical values.

Each of the files containing expression values should start with two integers, giving the number of genes G and the number of samples S_p, respectively. Following this, the expression values should be given as a $G \times S_p$ matrix. The order of the genes must be the same in all P files with expression values.

Each of the files containing clinical values should start with an integer, giving the number of samples, S_p. Following this, the S_q clinical values should be given as a vector of zeros and ones.

4.2 Parameter file

This file should have six lines. In each line a specified number of values should be given. Additional text in a line is just ignored. Thus, in each line comments may be added after the parameter values. To specify that a default value should be used for a parameter value, a '=' or a string with '=' as the first character can be given instead of a numerical value. If the file contains less than five lines, default values are used for the missing lines. If the file contains more than five lines, the extra lines are ignored.

The parameters given in each of the five lines are (default values are given in parenthesis):

1. Four values: $< \alpha_a(1.0) > < \beta_a(1.0) > < p_a^0(0.0) > < p_a^1(0.0) >$
2. Four values: $< \alpha_b(1.0) > < \beta_b(1.0) > < p_b^0(0.0) > < p_b^1(0.0) >$
3. One value: $< \nu_r(P+1) >$
4. One value: $< \nu_\rho(P+1) >$
5. Two values: $< \alpha_\xi(1.0) > < \beta_\xi(1.0) >$
6. One value: $< c^2_{\rm Max}(1.0) >$

4.3 Initial data file

This file should have $16 + P$ lines. Except lines 7 and 8, each line contains one value, the initial value for one parameter. Lines 7 and 8 can contain up to $P(P-1)/2$ values. Additional text in a line is again ignored and can be used for comments. A '=' or a string with '=' as the first character can again be used to specify that default values should be used (for lines 7 and 8 the '=' option can only be used when all values should take default values). If the file contains less than $16 + P$ lines, default values are used for the missing lines. If the file contains more than $16 + P$ lines, the extra lines are ignored.

Note that a parameter may be fixed be specifying its value in this file and specifying zero for the corresponding number of updates in the update file.

The initial values specified in each of the $16 + P$ lines are (default values are random and not specified here):

1. $< \nu_{gp} >$. If a value is specified, the same value is used for all components.
2. $< \Delta_{gp} >$. If a value is specified, the same value is used for all components.
3. $< a_p >$. If a value is specified, the same value is used for all components.
4. $< b_p >$. If a value is specified, the same value is used for all components.
5. $< c^2 >$.
6. $< \gamma^2 >$.

7. $< r_{pq} >$. Up to $P(P-1)/2$ values can be specified. Values are specified in the order $r_{12}, r_{13}, \ldots, r_{1P}, r_{23}, \ldots, r_{2P}$ and so on until $r_{P-1,P}$. If fewer than $P(P-1)/2$ values are specified, the last value is reused for all the remaining variables. If more than $P(P-1)/2$ values are specified, the extra values are ignored. The specified values must, together with 1's on the diagonal, form a symmetric positive definite matrix.

8. $< \rho_{pq} >$. Up to $P(P-1)/2$ values can be specified. Values are specified in the order $\rho_{12}, \rho_{13}, \ldots, \rho_{1P}, \rho_{23}, \ldots, \rho_{2P}$ and so on until $\rho_{P-1,P}$. If fewer than $P(P-1)/2$ values are specified, the last value is reused for all the remaining variables. If more than $P(P-1)/2$ values are specified, the extra values are ignored. The specified values must, together with 1's on the diagonal, form a symmetric positive definite matrix.

9. $< \delta_{gp} >$. If a value is specified, the same value is used for all components.

10. $< \xi_p >$. If a value is specified, the same value is used for all components.

11. $< \sigma^2_{gp} >$. If a value is specified, the same value is used for all components.

12. $< t_p >$. If a value is specified, the same value is used for all components.

13. $< l_p >$. If a value is specified, the same value is used for all components.

14. $< \phi_{gp} >$. If a value is specified, the same value is used for all components.

15. $< \theta_p >$. If a value is specified, the same value is used for all components.

16. $< \lambda_p >$. If a value is specified, the same value is used for all components.

17. $< \tau^2_1 >$.
    $\vdots$

16+P. $< \tau^2_P >$.

If the product of the values given for $\tau^2_1, \ldots, \tau^2_P$ differ from unity, these values are corrected by multiplying them all with the same value.

5 Update file

This file should have 17 lines, where each line specify one or two values. In line i, the first number gives the number of updates of type i (as defined in Section 3) in one iteration. If appropriate, the second number in line i specify the tuning parameter for update type i. Additional text in a line is again ignored and can be used for comments. A '=' or a string with '=' as the first character can again be used to specify that a default value should be used. If the file contains less than 17 lines, default values are used for the missing lines. If the file contains more than 17 lines, the extra lines are ignored.

Note that using a number larger than one for Gibbs updates has no effect for the convergence.

6 Output file

This file should have 22 lines. Each line contains one value. Additional text in a line is again ignored and can be used for comments. A '=' or a string with '=' as the first character can again be used to specify that default values should be used. If the file contains less than 20 lines, default values are used for the missing lines. If the file contains more than 20 lines, the extra lines are ignored.

The parameters given in the 22 lines are:

1. The number of Metropolis–Hastings iterations between each output to log files. Default value is 1.

2. File name of an output file with potential (minus log density) values. Default is no output. A total of nineteen values are written to the file. The first value is the (total) posterior potential, the second value is the potential associated with the likelihood. Thereafter follow potentials associated with each of the 17 variables specified in Section 4.3 and in the same order as given there.

3. File name of an output file with acceptance rates of Metropolis–Hastings proposals. Default is no output. A total of 17 values are written to the file, one value for each of the 17 proposal types specified in Section 3, and in the same order as in this section.

4. File name of an output file with ν_{gp} values. Default is no output. A total of GP values is written to file and the order is ν_{11}, ..., ν_{1P}, ν_{21}, ..., ν_{2P} and so on.

5. File name of an output file with $\delta_{gp} \cdot \Delta_{gp}$ values. Default is no output. A total of GP values is written to file and the order is as for ν_{gp} above.

6. File name of an output file with a_p values. Default is no output. A total of P values is written to file and the order is $a_1, \ldots, a_P$.

7. File name of an output file with b_p values. Default is no output. A total of P values is written to file and the order is as for a_p above.

8. File name of an output file with c^2 value. Default is no output.

9. File name of an output file with γ^2 value. Default is no output.

10. File name of an output file with r_{pq} values. Default is no output. A total of $P(P-1)/2$ values is written to file and the order is $r_{12}, \ldots, r_{1P}, r_{23}, \ldots, r_{2P}, r_{34}$ and so on.

11. File name of an output file with ρ_{pq} values. Default is no output. A total of $P(P-1)/2$ values is written to file and the order is as for r_{pq} above.

12. File name of an output file with δ_{gp} values. Default is no output. A total of GP values is written to file and the order is as for ν_{gp} above.

13. File name of an output file with ξ_p (prior model A) or ξ (prior model B) value. Default is no output. A total of P values are written to file and the order is $\xi_1, \ldots, \xi_P$. If prior model B is used the same value ξ is written p times.

14. File name of an output file with σ^2_{gp} values. Default is no output. A total of GP values is written to file and the order is as for ν_{gp} above.

15. File name of an output file with t_p values. Default is no output. A total of P values is written to file and the order is as for a_p above.

16. File name of an output file with l_p values. Default is no output. A total of P values is written to file and the order is as for a_p above.

17. File name of an output file with ϕ_{gp} values. Default is no output. A total of GP values is written to file and the order is as for ν_{gp} above.

18. File name of an output file with θ_p values. Default is no output. A total of P values is written to file and the order is as for a_p above.

19. File name of an output file with λ_p values. Default is no output. A total of P values is written to file and the order is as for a_p above.

20. File name of an output file with τ^2_p values. Default is no output. A total of P values is written to file and the order is as for a_p above.

21. File name of an output file with $\mathrm{P}(\delta_{gp} = 1|\ldots)$ values. Default is no output. A total of GP values is written to file and the order is as for ν_{gp} above.

22. File name of an output file where the running averages of differential expression, concordant differential expression and discordant differential expression is written. This file is no longer used. No output.

7 Call to diffExpressed_main.so from R

The diffExpressed_main.so takes a total of 54 arguments. Note that in the C++-code no checking of the input values is done, so make sure all input values are legal and all vectors are sufficiently long. The arguments are

1. Seed value. Input/Output. At input a positive integer. At output a new positive integer that can be used as seed if one wants to continue the same run for more iterations.

2. Number of Metropolis–Hastings iterations. Input only. Legal values: positive integers.

3. Indicator for printing iteration numbers to screen. Input only. Legal values: 0 and 1.

4. P: Number of studies. Input only.

5. G: Number of genes. Input only.

6. Prior model indicator, must be 0 or 1. If 0 is set, prior model A for δ_{gp} is used. Otherwise prior model B is used.

7. S: A vector of P values, giving the number of samples in each study. Input only.

8. x: A vector of $PG\sum_{p=1}^{P} S_p$ expression values. Input only. The order of the values is $x_{111}, x_{121}, \ldots, x_{1,S_1,1}, x_{2,1,1}, x_{2,2,1}, \ldots, x_{2,S_1,1}, \ldots, x_{G,S_1,1}, \ldots, x_{112}, x_{122}$ and so on.

9. ψ: A vector of $P\sum_{p=1}^{P} S_p$ clinical values. Input only. The order of the values is $\psi_{11}, \psi_{21}, \ldots, \psi_{S_1,1}, \psi_{12}, \psi_{22}, \ldots, \psi_{S_2,2}, \ldots, \psi_{S_P,P}$.

10. Indicator for using specified input values. Input only. If 1 is given, the simulation will used initial values as specified in the following 17 parameters. If 0 is given the initial values will be sampled randomly.

11. ν: Initial values for ν_{gp}. Input/Output. The order of the values is $\nu_{11}, \nu 21, \ldots, \nu_{G1}, \nu_{12}$, $\nu_{22}, \ldots, \nu_{G2}$ and so on. At output the vector contains the simulated values after the last Metropolis–Hastings iteration. Note that the vector must be of correct size (GP) even if argument 10 specifies that the input values should not be used as initial values.

12. Δ: Initial values for Δ_{gp}. Input/Output. The order is as for ν above. At output the vector contains the simulated values after the last Metropolis–Hastings iteration. Again the size of the vector must be of the correct size (GP).

13. a: Initial values for a_p. Input/Output. The order is $a_1, \ldots, a_P$. At output the vector contains the simulated values after the last Metropolis–Hastings iteration. Again the size of the vector must be of the correct size (P).

14. b: Initial values for b_p. Input/Output. The order is as for a above. At output the vector contains the simulated values after the last Metropolis–Hastings iteration. Again the size of the vector must be of the correct size (P).

15. c^2: Initial values for c^2. Input/Output. At output it contains the simulated value after the last Metropolis–Hastings iteration. Again the size of the vector must be of the correct size (1).

16. γ^2: Initial values for γ^2. Input/Output. At output it contains the simulated value after the last Metropolis–Hastings iteration. Again the size of the vector must be of the correct size (1).

17. r: Initial values for r_{pq}. Input/Output. The order is $r_{12}, r_{13}, \ldots, r_{1P}, r_{23}, r_{24}, \ldots, r_{2P}$, $\ldots, r_{P-1,P}$. At output the vector contains the simulated values after the last Metropolis–Hastings iteration. Again the size of the vector must be of the correct size ($P(P-1)/2$).

18. ρ: Initial values for ρ_{pq}. Input/Output. The order is as for r above. At output the vector contains the simulated values after the last Metropolis–Hastings iteration. Again the size of the vector must be of the correct size ($P(P-1)/2$).

19. δ: Initial values for δ_{gp}. Input/Output. The order is as for ν_{gp} above. At output the vector contains the simulated values after the last Metropolis–Hastings iteration. Again the size of the vector must be of the correct size (GP).

20. ξ: Initial values for ξ_p. Input/Output. At output it contains the simulated value after the last Metropolis–Hastings iteration. Again the size of the vector must be of the correct size (P).

21. σ_2: Initial values for σ^2_{gp}. Input/Output. The order is as for ν above. At output the vector contains the simulated values after the last Metropolis–Hastings iteration. Again the size of the vector must be of the correct size (GP).

22. t: Initial values for t_p. Input/Output. The order is as for a above. At output the vector contains the simulated values after the last Metropolis–Hastings iteration. Again the size of the vector must be of the correct size (P).

23. l: Initial values for l_p. Input/Output. The order is as for a above. At output the vector contains the simulated values after the last Metropolis–Hastings iteration. Again the size of the vector must be of the correct size (P).

24. ϕ: Initial values for ϕ_{gp}. Input/Output. The order is as for ν above. At output the vector contains the simulated values after the last Metropolis–Hastings iteration. Again the size of the vector must be of the correct size (GP).

25. θ: Initial values for θ_p. Input/Output. The order is as for a above. At output the vector contains the simulated values after the last Metropolis–Hastings iteration. Again the size of the vector must be of the correct size (P).

26. λ: Initial values for λ_p. Input/Output. The order is as for a above. At output the vector contains the simulated values after the last Metropolis–Hastings iteration. Again the size of the vector must be of the correct size (P).

27. τ^2: Initial values for τ^2_p. Input/Output. The order is as for a above. At output the vector contains the simulated values after the last Metropolis–Hastings iteration. Again the size of the vector must be of the correct size (P).

28. Hyper-parameter values: A vector of 12 values. Input only. The order is the same as specified in Section 4.2.

29. Number of updates of each type in one Metropolis–Hastings iterations: A vector of 17 positive integers. Input only. Element i specifies number of updates of type i, see Section 3.

30. ε: A vector of 17 positive real numbers. Input only. Element i specifies the ε for update type i. If update type i doesn't include any tuning parameter ε, the value is just ignored.

31. A vector of 22 elements. Input only. The first element specifies the number of Metropolis–Hastings iterations between the storing of simulated values. Elements 2 to 22 are indicators specifying whether or not to store a specific variable type. The type of variable types and the order used is identical to what is used in Section 6. The stored values are either written to files or stored in memory, see the next parameter.

32. Write to file indicator. Input only. If a non-zero value is given, simulated values are written to files, otherwise the simulated values are stored in memory (which may require a lot of memory). See also 55.

33. Name of directory for where to write files of simulated values. Input only. One file is produced for each of the 17 variables, the file names are predefined. If parameter 32 is 0, this parameter is ignored.

34. Simulated potential values (in one long vector). Output only. The same potential values as previously written to file, and in the same order. Note that at input the vector size must be large enough to store all the values produced. If parameter 32 is set to 1 or the corresponding element in parameter 31 is zero, this variable is not used.

35. As item 34, but Metropolis–Hastings acceptance rates.

36. As item 34, but simulated ν-values.

37. As item 34, but simulated Δ-values.

38. As item 34, but simulated a-values.

39. As item 34, but simulated b-values.

40. As item 34, but simulated c^2-values.

41. As item 34, but simulated γ^2-values.

42. As item 34, but simulated r-values.

43. As item 34, but simulated ρ-values.

44. As item 34, but simulated δ-values.

45. As item 34, but simulated ξ-values.

46. As item 34, but simulated σ^2-values.

47. As item 34, but simulated t-values.

48. As item 34, but simulated l-values.

49. As item 34, but simulated ϕ-values.

50. As item 34, but simulated θ-values.

51. As item 34, but simulated λ-values.

52. As item 34, but simulated τ^2-values.

53. As item 34, but simulated $\mathrm{P}(\delta_{gp} = 1 | \ldots)$-values.

54. Averages of differential expression, concordant differential expression and discordant differential expression (all in one long vector). Output only. Unlike the previous output arguments, only the final values are given as output. Thus, the vector size must be $3G$. The order of the output values is first all averages for differential expression, followed by all averages for concordant differential expression and then all running averages for all discordant differential expression. If the corresponding element in parameter 31 is zero, this variable is not used. If parameter 32 is set to 1 this variable is only used if parameter 55 is set to 1.

55. Indicator for using parameter 54.