A New Summarization Method for Affymetrix Probe Level Data
SUPPLEMENTARY INFORMATION

Sepp Hochreiter, Djork-Arne Clevert, and Klaus Obermayer
Department of Electrical Engineering and Computer Science
Technische Universität Berlin
10587 Berlin, Germany
{hochreit,okko,oby}@cs.tu-berlin.de

7. October 2005

Contents

1 Maximum A Posteriori Factor Analysis

The model parameters are estimated by the maximum a posteriori estimation (e.g. Gauvain and Lee, 1994) which extends the maximum likelihood estimation by introducing a prior for the parameters (e.g. DeGroot, 1970). We apply the iterative Expectation-Maximization (EM) optimization technique Dempster et al. (1977) analog to the maximum likelihood approach (e.g. Rubin and Thayer, 1982; Everitt, 1984).

We are given the data $\{\boldsymbol{x}\} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$ which is already normalized to mean zero (by subtracting the mean $\boldsymbol{\mu}$ from the data) and assume one factor. The model is

$$\boldsymbol{x} \;=\; \boldsymbol{\lambda} z \;+\; \boldsymbol{\epsilon} \;, \tag{1}$$

where

$$z \;\sim\; \mathcal{N}\left(0, 1\right) \quad \text{and} \quad \boldsymbol{\epsilon} \;\sim\; \mathcal{N}\left(\mathbf{0}, \boldsymbol{\Psi}\right) \;. \tag{2}$$

Here $\boldsymbol{x}, \boldsymbol{\epsilon}, \boldsymbol{\lambda} \in \mathbb{R}^n$, $z \in \mathbb{R}$, $\boldsymbol{\Psi}$ is a diagonal matrix from $\mathbb{R}^{n \times n}$, and $\boldsymbol{\epsilon}$ and z are independent. It follows that

$$\boldsymbol{x} \mid z \;\;\sim\;\; \mathcal{N}\left(\boldsymbol{\lambda} z, \boldsymbol{\Psi}\right) \;. \tag{3}$$

The posterior $p(\boldsymbol{\lambda}, \boldsymbol{\Psi} \mid \{\boldsymbol{x}\})$ is proportional to the product between the likelihood $p(\{\boldsymbol{x}\} \mid \boldsymbol{\lambda}, \boldsymbol{\Psi})$ and the prior $p(\boldsymbol{\lambda})$:

$$p(\boldsymbol{\lambda}, \boldsymbol{\Psi} \mid \{\boldsymbol{x}\}) \;\propto\; p(\{\boldsymbol{x}\} \mid \boldsymbol{\lambda}, \boldsymbol{\Psi})\; p(\boldsymbol{\lambda}) \;, \tag{4}$$

therefore up to a constant independent of the parameters the log-posterior is

$$\log\left(p(\boldsymbol{\lambda}, \boldsymbol{\Psi} \mid \{\boldsymbol{x}\})\right) \;=\; \log\left(p(\{\boldsymbol{x}\} \mid \boldsymbol{\lambda}, \boldsymbol{\Psi})\right) \;+\; \log\left(p(\boldsymbol{\lambda})\right) \;. \tag{5}$$

The prior on λ_j is a rectified Gaussian $\mathcal{N}_{\text{rect}}\left(\mu_\lambda, \sigma_\lambda\right)$:

$$y_j \;\sim\; \mathcal{N}\left(\mu_\lambda, \sigma_\lambda\right) \tag{6}$$

$$\lambda_j \;=\; \max\{y_j, 0\} \;. \tag{7}$$

The rectification can be expressed by a weighted sum of a Gaussian $\mathcal{N}_1\left(0, \sigma_{\text{rect}}\right)$ with zero mean and variance $\sigma_{\text{rect}} \;\rightarrow\; 0$ and a Gaussian density $\mathcal{N}_2\left(\mu_\lambda, \sigma_\lambda\right)$ cut at zero. Note, that

$$\begin{aligned} \frac{\partial \log\left(w_1 \mathcal{N}_1 \;+\; w_2 \mathcal{N}_2\right)}{\partial \lambda_j} \;&=\; \frac{1}{w_1 \mathcal{N}_1 \;+\; w_2 \mathcal{N}_2} \left(w_1 \mathcal{N}_1 \;+\; w_2 \mathcal{N}_2\right) \\ \left(-\, \sigma_\lambda^{-2}\left(\lambda_j \;-\; \mu_\lambda\right) -\; \sigma_{\text{rect}}^{-2} \lambda_j\right) \;&=_{\lambda_j \neq 0}\; -\, \sigma_\lambda^{-2}\left(\lambda_j \;-\; \mu_\lambda\right) \;. \end{aligned} \tag{8}$$

Therefore we only consider the the logarithm of the positive part of the prior $\mathcal{N}_{\text{rect}}\left(\mu_\lambda \mathbf{1}, \sigma_\lambda \boldsymbol{I}\right)$, $\forall_j : \lambda_j > \mu_\lambda$, which is up to a constant which is independent of $\boldsymbol{\lambda}$:

$$-\, \frac{1}{2} \sigma_\lambda^{-2} \left(\boldsymbol{\lambda} \;-\; \mu_\lambda \mathbf{1}\right)^T \left(\boldsymbol{\lambda} \;-\; \mu_\lambda \mathbf{1}\right) \;. \tag{9}$$

We will now consider the likelihood of the data. Let denote E the expectation of the data (i.e. the factor distribution and the noise distribution is combined):

$$
\begin{aligned}
&\mathrm{E}(\boldsymbol{x}) \ = \ \mathrm{E}(\boldsymbol{\lambda} z \ + \ \boldsymbol{\epsilon}) \ = \ \boldsymbol{\lambda}\mathrm{E}(z) \ + \ \mathrm{E}(\boldsymbol{\epsilon}) \ = \ \mathbf{0} \ , \qquad (10)\\
&\mathrm{E}\left(\boldsymbol{x}\ \boldsymbol{x}^T\right) \ = \ \mathrm{E}\left((\boldsymbol{\lambda} z \ + \ \boldsymbol{\epsilon})(\boldsymbol{\lambda} z \ + \ \boldsymbol{\epsilon})^T\right) \ = \\
&\boldsymbol{\lambda}\mathrm{E}\left(z^2\right)\boldsymbol{\lambda}^T \ + \ \boldsymbol{\lambda}\mathrm{E}\left(z\right)\mathrm{E}\left(\boldsymbol{\epsilon}^T\right) \ \mathrm{E}\left(z\right)\mathrm{E}\left(\boldsymbol{\epsilon}\right) \ \boldsymbol{\lambda}^T \ + \ \mathrm{E}\left(\boldsymbol{\epsilon}\ \boldsymbol{\epsilon}^T\right) \ = \\
&\boldsymbol{\lambda}\ \boldsymbol{\lambda}^T \ + \ \boldsymbol{\Psi} \ .
\end{aligned}
$$

Therefore, the marginal distribution for $\boldsymbol{x}$ is

$$
\boldsymbol{x} \ \sim \ \mathcal{N}\left(\mathbf{0}\ , \ \boldsymbol{\lambda}\boldsymbol{\lambda}^T \ + \ \boldsymbol{\Psi}\right) \ . \qquad (11)
$$

The log-likelihood of the data $\{\boldsymbol{x}_i\}$ under the model $(\boldsymbol{\lambda}, \boldsymbol{\Psi})$ is

$$
\log \prod_{i=1}^{N} (2\pi)^{-n/2} \left|\boldsymbol{\lambda}\boldsymbol{\lambda}^T \ + \ \boldsymbol{\Psi}\right|^{-1/2} \ \exp\left(-\frac{1}{2}\left(\boldsymbol{x}_i^T \left(\boldsymbol{\lambda}\boldsymbol{\lambda}^T \ + \ \boldsymbol{\Psi}\right)^{-1} \boldsymbol{x}_i\right)\right) \ , \qquad (12)
$$

where $|.|$ denotes the absolute value of the determinant of a matrix. The Maximum a Posteriori approach maximizes the sum of the log-likelihood of the data and the log-prior. But even to maximize only the likelihood is difficult because no closed form of directly maximizing the likelihood with respect to the parameters is known. Therefore, the maximum likelihood approach to factor analysis is in most cases (e.g. Jöreskog, 1967) based on the Expectation-Maximization (EM) optimization technique Dempster et al. (1977). We will also apply the EM for our Maximum a Posteriori approach.

Using

$$
Q_i(z_i) \ = \ p\left(z_i \mid \boldsymbol{x}_i; \boldsymbol{\lambda}, \boldsymbol{\Psi}\right) \qquad (13)
$$

then

$$
\begin{aligned}
z_i \mid \boldsymbol{x}_i \ &\sim \ \mathcal{N}\left(\mu_{z_i|\boldsymbol{x}_i}, \sigma^2_{z_i|\boldsymbol{x}_i}\right) \qquad (14)\\
\mu_{z_i|\boldsymbol{x}_i} \ &= \ (\boldsymbol{x}_i)^T \ \left(\boldsymbol{\lambda}\ \boldsymbol{\lambda}^T \ + \ \boldsymbol{\Psi}\right)^{-1} \boldsymbol{\lambda}\\
\sigma^2_{z_i|\boldsymbol{x}_i} \ &= \ 1 \ - \ \boldsymbol{\lambda}^T \ \left(\boldsymbol{\lambda}\ \boldsymbol{\lambda}^T \ + \ \boldsymbol{\Psi}\right)^{-1} \boldsymbol{\lambda} \ ,
\end{aligned}
$$

where we used the fact that

$$
\begin{aligned}
&\boldsymbol{v} \ \sim \ \mathcal{N}\left(\boldsymbol{\mu}_v, \Sigma_{vv}\right) \ , \ \boldsymbol{u} \ \sim \ \mathcal{N}\left(\boldsymbol{\mu}_u, \Sigma_{uu}\right) \ , \qquad (15)\\
&\Sigma_{uv} \ = \ \mathrm{Covar}(\boldsymbol{u}, \boldsymbol{v}) \text{ and } \ \Sigma_{vu} \ = \ \mathrm{Covar}(\boldsymbol{v}, \boldsymbol{u}) :\\
&\boldsymbol{v} \mid \boldsymbol{u} \ \sim \ \mathcal{N}\left(\boldsymbol{\mu}_v \ + \ \Sigma_{vu}\Sigma_{uu}^{-1}\left(\boldsymbol{u} \ - \ \boldsymbol{\mu}_u\right) \ , \ \Sigma_{vv} \ + \ \Sigma_{vu}\Sigma_{uu}^{-1}\Sigma_{uv}\right)
\end{aligned}
$$

and

$$
\mathrm{E}(z\boldsymbol{x}) \ = \ \boldsymbol{\lambda}\mathrm{E}(z^2) \ = \ \boldsymbol{\lambda} \ . \qquad (16)
$$

We obtain

$$
Q_i(z_i) \ = \ (2\pi)^{-1/2} \, \sigma^{-1}_{z_i|\boldsymbol{x}_i} \exp\left(-\frac{1}{2\ \sigma^2_{z_i|\boldsymbol{x}_i}} \left(z_i \ - \ \mu_{z_i|\boldsymbol{x}_i}\right)^2\right) \ . \qquad (17)
$$

The EM algorithm for maximum posterior maximizes in the M-step a lower bound for the posterior:

$$\begin{aligned}
&\log \left(p(\boldsymbol{x}_i \mid \boldsymbol{\lambda}, \boldsymbol{\Psi})\ p(\boldsymbol{\lambda})\right) \ = \\
&\log \left(\int_{-\infty}^{+\infty} \frac{Q_i(z_i)\ p(\boldsymbol{x}_i, z_i \mid \boldsymbol{\lambda}, \boldsymbol{\Psi})\ p(\boldsymbol{\lambda})}{Q_i(z_i)} dz_i \right) \ \geq \\
&\int_{-\infty}^{+\infty} Q_i(z_i)\ \log \left(\frac{p(\boldsymbol{x}_i, z_i \mid \boldsymbol{\lambda}, \boldsymbol{\Psi})\ p(\boldsymbol{\lambda})}{Q_i(z_i)} \right) dz_i .
\end{aligned} \tag{18}$$

Using the expectation

$$\mathrm{E}_{z_i \mid \boldsymbol{x}_i} \left(f(z_i)\right) \ = \ \int_{-\infty}^{\infty} Q_i(z_i)\ f(z_i)\ dz_i \tag{19}$$

and neglecting all terms which are independent of $\boldsymbol{\lambda}$ and $\boldsymbol{\Psi}$, the M-step requires to maximize

$$\begin{aligned}
&O \ = \ - \ \frac{1}{2} \sigma_{\lambda}^{-2} \left(\boldsymbol{\lambda} \ - \ \mu_{\lambda} \mathbf{1}\right)^T \left(\boldsymbol{\lambda} \ - \ \mu_{\lambda} \mathbf{1}\right) \ + \ \frac{nN}{2} \log\left(2\pi\right) \\
&- \frac{N}{2} \log \left|\boldsymbol{\Psi}\right| \ - \ \frac{1}{2} \sum_{i=1}^{N} \mathrm{E}_{z_i \mid \boldsymbol{x}_i} \left(\left(\boldsymbol{x}_i \ - \ \boldsymbol{\lambda} z_i\right)^T \boldsymbol{\Psi}^{-1} \left(\boldsymbol{x}_i \ - \ \boldsymbol{\lambda} z_i\right) \right) \ .
\end{aligned} \tag{20}$$

The optimality criteria are

$$\begin{aligned}
&\frac{1}{N} \nabla_{\boldsymbol{\lambda}} O \ = \ - \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{\Psi}^{-1}\ \boldsymbol{\lambda}\ \mathrm{E}_{z_i \mid \boldsymbol{x}_i} \left(z_i^2\right) \ + \\
&\frac{1}{N} \sum_{i=1}^{N} \boldsymbol{\Psi}^{-1}\ \boldsymbol{x}_i\ \mathrm{E}_{z_i \mid \boldsymbol{x}_i} \left(z_i\right) \ - \ \frac{1}{N} \sigma_{\lambda}^{-2} \left(\boldsymbol{\lambda} \ - \ \mu_{\lambda} \mathbf{1}\right) \ = \ \mathbf{0}
\end{aligned} \tag{21}$$

and

$$\begin{aligned}
&\nabla_{\boldsymbol{\Psi}} O \ = \ - \frac{N}{2} \boldsymbol{\Psi}^{-1} \ + \\
&\frac{1}{2} \sum_{i=1}^{N} \mathrm{E}_{z_i \mid \boldsymbol{x}_i} \left(\boldsymbol{\Psi}^{-1} \left(\boldsymbol{x}_i \ - \ \boldsymbol{\lambda} z_i\right) \left(\boldsymbol{x}_i \ - \ \boldsymbol{\lambda} z_i\right)^T \boldsymbol{\Psi}^{-1} \right) \ = \ \mathbf{0} \ .
\end{aligned} \tag{22}$$

Solving above equations gives:

$$\lambda_j^{\text{Gauss}} = \left(\frac{1}{N} \sum_{i=1}^{N} x_{ij} \ \mathrm{E}_{z_i|\boldsymbol{x}_i} \left(z_i \right) \ + \ \frac{1}{N} \frac{\mu_\lambda \ \Psi_{jj}^{\text{new}}}{\sigma_\lambda^2} \right) \tag{23}$$
$$\left(\frac{1}{N} \sum_{i=1}^{N} \mathrm{E}_{z_i|\boldsymbol{x}_i} \left(z_i^2 \right) \ + \ \frac{1}{N} \frac{\Psi_{jj}^{\text{new}}}{\sigma_\lambda^2} \right)^{-1} ,$$
$$\lambda_j^{\text{new}} = \begin{cases} \lambda_j^{\text{Gauss}} & \text{for} \quad \lambda_j^{\text{Gauss}} \ > \ 0 \\ 0 & \text{for} \quad \lambda_j^{\text{Gauss}} \ \leq \ 0 \end{cases} ,$$
$$\boldsymbol{\Psi}^{\text{new}} = \mathrm{diag} \left(\frac{1}{N} \sum_{i=1}^{N} \mathrm{E}_{z_i|\boldsymbol{x}_i} \left(\left(\boldsymbol{x}_i \ - \ \boldsymbol{\lambda}^{\text{new}} z_i \right) \left(\boldsymbol{x}_i \ - \ \boldsymbol{\lambda}^{\text{new}} z_i \right)^T \right) \right) = \tag{24}$$
$$\mathrm{diag} \left(\frac{1}{N} \sum_{i=1}^{N} \boldsymbol{x}_i \boldsymbol{x}_i^T \ - \ \frac{1}{N} \sum_{i=1}^{N} \mathrm{E}_{z_i|\boldsymbol{x}_i} \left(z_i \right) \boldsymbol{x}_i \left(\boldsymbol{\lambda}^{\text{new}} \right)^T \ - \right.$$
$$\left. \frac{1}{N} \sum_{i=1}^{N} \mathrm{E}_{z_i|\boldsymbol{x}_i} \left(z_i \right) \boldsymbol{\lambda}^{\text{new}} \boldsymbol{x}_i^T \ + \ \frac{1}{N} \sum_{i=1}^{N} \mathrm{E}_{z_i|\boldsymbol{x}_i} \left(z_i^2 \right) \boldsymbol{\lambda}^{\text{new}} \left(\boldsymbol{\lambda}^{\text{new}} \right)^T \right) ,$$

where "diag" makes a diagonal matrix from a matrix by setting all non-diagonal elements to zero and "Gauss" denotes the solution for Gaussian prior without rectification. The backprojection to zero for $\lambda_k < 0$ is due to the prior. Note, that backprojecting negative λ_j to zero in the iterative algorithm only approximates the prior.

From eq. (23) we obtain for $\lambda_j^{\text{new}} \ > \ 0$

$$\lambda_j^{\text{new}} \frac{1}{N} \sum_{i=1}^{N} \mathrm{E}_{z_i|\boldsymbol{x}_i} \left(z_i^2 \right) \ = \tag{25}$$
$$- \ \lambda_j^{\text{new}} \frac{1}{N} \frac{\Psi_{jj}^{\text{new}}}{\sigma_\lambda^2} \ + \ \frac{1}{N} \sum_{i=1}^{N} x_{ij} \ \mathrm{E}_{z_i|\boldsymbol{x}_i} \left(z_i \right) \ + \ \frac{1}{N} \frac{\mu_\lambda \ \Psi_{jj}^{\text{new}}}{\sigma_\lambda^2}$$

Inserting eq. (25) into the last term of eq. (24) gives (note, that only the diagonal elements are relevant):

$$\Psi_{jj}^{\text{new}} = \left[\mathrm{diagvect} \left(\frac{1}{N} \sum_{i=1}^{N} \boldsymbol{x}_i \boldsymbol{x}_i^T \right) \right]_j \ - \ \lambda_j^{\text{new}} \left[\frac{1}{N} \sum_{i=1}^{N} \mathrm{E}_{z_i|\boldsymbol{x}_i} \left(z_i \right) \boldsymbol{x}_i \right]_j \ +$$
$$\frac{1}{N} \frac{\Psi_{jj}^{\text{new}}}{\sigma_\lambda^2} \lambda_j^{\text{new}} \left(\mu_\lambda \ - \ \lambda_j^{\text{new}} \right) , \tag{26}$$

where "diagvect(.)" gives the vector consisting of the diagonal elements of a matrix. Note, that that this reformulation also holds for $\lambda_j^{\text{new}} \ = \ 0$, i.e. holds also for rectification.

This leads to following EM updates:

$$\textbf{E-step:} \tag{27}$$
$$\begin{aligned}
\mathrm{E}_{z_i|\boldsymbol{x}_i}\left(z_i\right) &= \mu_{z_i|\boldsymbol{x}_i} \\
\mathrm{E}_{z_i|\boldsymbol{x}_i}\left(z_i^2\right) &= \mu_{z_i|\boldsymbol{x}_i}^2 \ + \ \sigma_{z_i|\boldsymbol{x}_i}^2
\end{aligned}$$

$$\textbf{M-step:} \tag{28}$$
$$\begin{aligned}
\lambda_j^{\mathrm{Gauss}} &= \left(\frac{1}{N}\sum_{i=1}^{N} x_{ij} \ \mathrm{E}_{z_i|\boldsymbol{x}_i}\left(z_i\right) \ + \ \frac{1}{N}\frac{\mu_\lambda \ \Psi_{jj}^{\mathrm{new}}}{\sigma_\lambda^2}\right) \\
& \quad \left(\frac{1}{N}\sum_{i=1}^{N} \mathrm{E}_{z_i|\boldsymbol{x}_i}\left(z_i^2\right) \ + \ \frac{1}{N}\frac{\Psi_{jj}^{\mathrm{new}}}{\sigma_\lambda^2}\right)^{-1} , \\
\lambda_j^{\mathrm{new}} &= \begin{cases} \lambda_j^{\mathrm{Gauss}} & \text{for} \quad \lambda_j^{\mathrm{Gauss}} \ > \ 0 \\ 0 & \text{for} \quad \lambda_j^{\mathrm{Gauss}} \ \leq \ 0 \end{cases} , \\
\Psi_{jj}^{\mathrm{new}} &= \left[\mathrm{diagvect}\left(\frac{1}{N}\sum_{i=1}^{N} \boldsymbol{x}_i \boldsymbol{x}_i^T\right)\right]_j \ - \ \lambda_j^{\mathrm{new}} \left[\frac{1}{N}\sum_{i=1}^{N} \mathrm{E}_{z_i|\boldsymbol{x}_i}\left(z_i\right) \boldsymbol{x}_i\right]_j \ + \\
& \quad \frac{1}{N}\frac{\Psi_{jj}^{\mathrm{new}}}{\sigma_\lambda^2} \lambda_j^{\mathrm{new}} \left(\mu_\lambda \ - \ \lambda_j^{\mathrm{new}}\right) \ . \qquad (29)
\end{aligned}$$

The actual update must be done by solving the equations in the M-step for $\boldsymbol{\Psi}^{\mathrm{new}}$ and $\boldsymbol{\lambda}^{\mathrm{new}}$ which is complicated due to the nonlinear equations. However in the next appendix a more efficient approximation to the solution of the M-step is described.

To allow the factor loading $\boldsymbol{\lambda}$ in the range of the data variance we set

$$\sigma_\lambda^2 \ = \ \rho \ \frac{1}{n}\sum_{j=1}^{n} \mathrm{Var}\left(\boldsymbol{x}\right) \ . \tag{30}$$

In conclusion, the algorithm possesses two hyperparameters μ_λ and ρ, where μ_λ determines whether false positives or false negatives should be reduced and ρ determines the influence of the prior.

2 Algorithm

The algorithm is available as an **R** package from `http://ni.cs.tu-berlin.de/~hochreit/software/affy/`.

2.1 Approximation

The M-step in its current form is not efficient because it must be solved for $\boldsymbol{\Psi}^{\mathrm{new}}$ and $\boldsymbol{\lambda}^{\mathrm{new}}$. Therefore we approximate Ψ_{jj}^{new} in the right hand side of the M-step in eq. (28) by Ψ_{jj}^{old} and also in the right hand side of eq. (29) because $\boldsymbol{\lambda}^{\text{new}}$ was used in to derive this equation. This approximation is justified because the original algorithms converges to $\Psi_{jj}^{\text{new}} = \Psi_{jj}^{\text{old}}$ and, therefore, the solution of the approximative algorithms is equal to the EM-solution. The question remains: does the approximative algorithm converge? It would not converge if the approximative M-step does not decrease the a posteriori but we never observed this in our simulations of many 100.000 cases.

2.2 Algebraic Reformulations / Efficiency

The matrix inversion lemma allows us to compute $\left(\boldsymbol{\lambda}\,\boldsymbol{\lambda}^T + \boldsymbol{\Psi}\right)^{-1}$ as follows:

$$\begin{aligned} \left(\boldsymbol{\lambda}\,\boldsymbol{\lambda}^T + \boldsymbol{\Psi}\right)^{-1} &= \boldsymbol{\Psi}^{-1} - \boldsymbol{\Psi}^{-1}\boldsymbol{\lambda}\left(1 + \boldsymbol{\lambda}^T\boldsymbol{\Psi}^{-1}\boldsymbol{\lambda}\right)^{-1}\boldsymbol{\lambda}^T\boldsymbol{\Psi}^{-1} = \\ & \boldsymbol{\Psi}^{-1} - \left(1 + \boldsymbol{\lambda}^T\boldsymbol{\Psi}^{-1}\boldsymbol{\lambda}\right)^{-1}\boldsymbol{\Psi}^{-1}\boldsymbol{\lambda}\boldsymbol{\lambda}^T\boldsymbol{\Psi}^{-1}\,. \end{aligned} \tag{31}$$

We obtain

$$\begin{aligned} &\left(\boldsymbol{\lambda}\,\boldsymbol{\lambda}^T + \boldsymbol{\Psi}\right)^{-1}\boldsymbol{\lambda} = \\ &\boldsymbol{\Psi}^{-1}\boldsymbol{\lambda} - \left(1 + \boldsymbol{\lambda}^T\boldsymbol{\Psi}^{-1}\boldsymbol{\lambda}\right)^{-1}\boldsymbol{\Psi}^{-1}\boldsymbol{\lambda}\boldsymbol{\lambda}^T\boldsymbol{\Psi}^{-1}\boldsymbol{\lambda} = \\ &\boldsymbol{\Psi}^{-1}\boldsymbol{\lambda}\left(1 - \left(1 + \boldsymbol{\lambda}^T\boldsymbol{\Psi}^{-1}\boldsymbol{\lambda}\right)^{-1}\boldsymbol{\lambda}^T\boldsymbol{\Psi}^{-1}\boldsymbol{\lambda}\right) = \\ &\boldsymbol{\Psi}^{-1}\boldsymbol{\lambda}\left(1 + \boldsymbol{\lambda}^T\boldsymbol{\Psi}^{-1}\boldsymbol{\lambda}\right)^{-1} \end{aligned} \tag{32}$$

Using eq. (32), we obtain new formulas for the eqs. (14)

$$\begin{aligned} \mu_{z_i|\boldsymbol{x}_i} &= (\boldsymbol{x}_i)^T\ \boldsymbol{\Psi}^{-1}\boldsymbol{\lambda}\left(1 + \boldsymbol{\lambda}^T\boldsymbol{\Psi}^{-1}\boldsymbol{\lambda}\right)^{-1} \\ \sigma^2_{z_i|\boldsymbol{x}_i} &= \left(1 + \boldsymbol{\lambda}^T\boldsymbol{\Psi}^{-1}\boldsymbol{\lambda}\right)^{-1}\,. \end{aligned} \tag{33}$$

Note that

$$\begin{aligned} &\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{x}_i\ \mathrm{E}_{z_i|\boldsymbol{x}_i}(z_i) = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{x}_i\boldsymbol{x}_i^T\left(\boldsymbol{\lambda}\,\boldsymbol{\lambda}^T + \boldsymbol{\Psi}\right)^{-1}\boldsymbol{\lambda} = \\ &\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{x}_i\boldsymbol{x}_i^T\boldsymbol{\Psi}^{-1}\boldsymbol{\lambda}\left(1 + \boldsymbol{\lambda}^T\boldsymbol{\Psi}^{-1}\boldsymbol{\lambda}\right)^{-1} \\ &\frac{1}{N}\sum_{i=1}^{N}\mathrm{E}_{z_i|\boldsymbol{x}_i}\left(z_i^2\right) = \boldsymbol{\lambda}^T\boldsymbol{\Psi}^{-1}\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{x}_i\boldsymbol{x}_i^T\boldsymbol{\Psi}^{-1}\boldsymbol{\lambda}\left(1 + \boldsymbol{\lambda}^T\boldsymbol{\Psi}^{-1}\boldsymbol{\lambda}\right)^{-2} + \\ &\left(1 + \boldsymbol{\lambda}^T\boldsymbol{\Psi}^{-1}\boldsymbol{\lambda}\right)^{-1}\,, \end{aligned} \tag{34}$$

therefore the covariance matrix $\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{x}_i\boldsymbol{x}_i^T$ can be computed once at the initialization of the algorithm and all sums $\sum_{i=1}^{N}$ vanish in the algorithm.

Above reformulations allow to efficiently compute the parameter updates using $\psi_j \; = \; \Psi_{jj}$ and the variables $\boldsymbol{\phi}, a, \boldsymbol{\eta}, \xi$ and $\boldsymbol{\zeta}$:

$$\forall_j : \phi_j \; = \; \frac{\lambda_j}{\psi_j} \tag{35}$$

$$a \; = \; \left(1 \; + \; \boldsymbol{\lambda}^T \boldsymbol{\phi}\right)^{-1} \tag{36}$$

$$\boldsymbol{\eta} \; = \; \boldsymbol{\phi} \; a \tag{37}$$

$$\boldsymbol{\zeta} \; = \; \mathrm{Covar}\left(\boldsymbol{x}\right) \; \boldsymbol{\eta} \quad \text{(first line of eq. (34))} \tag{38}$$

$$\xi \; = \; a \; + \; \boldsymbol{\eta}^T \; \boldsymbol{\zeta} \tag{39}$$

$$\forall_j : \lambda_j \; = \; \frac{\zeta_j \; + \; \psi_j \; \mu_\lambda \; \sigma_\lambda^{-2}}{\xi \; + \; \psi_j \; \sigma_\lambda^{-2}} \tag{40}$$

$$\forall_j : \; \psi_j \; = \; \mathrm{Covar}\left(\boldsymbol{x}\right)_{jj} \; - \; \zeta_j \lambda_j \; + \; \psi_j \; \sigma_\lambda^{-2} \; \lambda_j \left(\mu_\lambda - \lambda_j\right) \; . \tag{41}$$

2.3 Initialization

For the noise-free case factor analysis is equivalent to extracting the maximal eigenvalue of the covariance matrix:

$$\boldsymbol{x}_i \; = \; z_i \; \boldsymbol{\lambda} \tag{42}$$

$$\mathrm{Covar}\left(\boldsymbol{x}\right) \; = \; \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{x}_i \; \boldsymbol{x}_i^T \; = \; \boldsymbol{\lambda}\boldsymbol{\lambda}^T \frac{1}{N} \sum_{i=1}^{N} z_i^2 \; = \; \boldsymbol{\lambda}\boldsymbol{\lambda}^T \; ,$$

therefore $\boldsymbol{C} := \mathrm{Covar}\left(\boldsymbol{x}\right)$ has only one non-zero eigenvalue $\|\boldsymbol{\lambda}\|$ with corresponding eigenvector $\frac{\boldsymbol{\lambda}}{\|\boldsymbol{\lambda}\|}$. That is the motivation to initialize $\boldsymbol{\lambda}$ with the largest eigenvalue $e_{\max}$ of $\mathrm{Covar}\left(\boldsymbol{x}\right)$ multiplied with component-wise absolute value of the corresponding eigenvector $\boldsymbol{v}_{\max}$

$$\boldsymbol{\lambda}^{\mathrm{init}} \; = \; \left|e_{\max}\right| \; \left|v_j^{\max}\right| \; . \tag{43}$$

Because the data variance $\mathrm{Var}\left(x_j\right)$ of component j is explained through Ψ_{jj} and λ_j, we initialize $\boldsymbol{\Psi}$ as

$$\boldsymbol{\Psi}_{jj}^{\mathrm{init}} \; = \; \mathrm{Var}\left(x_j\right) \; . \tag{44}$$

2.4 Pseudo Code

Algorithmus 1 FARMS

Preprocessing: quantile or cyclic loess normalization (all genes and arrays)
Parameters: backscale factor f (default: 2.0 for quantile norm., 1.5 for cyclic loess); μ_λ (default: 0); prior variance factor ρ (default: $\frac{1}{8}$) , maximal cycles $\mathrm{cyc}_{\max}$ (default: 100); termination tolerance tol (default: 0.00001)

FORALL genes DO
compute maximal absolute eigenvalue $e_{\max}$ with eigenvector $\boldsymbol{v}_{\max}$ of $\boldsymbol{C}$
FOR j in 1:probes $\lambda_j^{\mathrm{init}} = |e_{\max}| \, \left|v_j^{\max}\right|$ **END FOR** {cf. eq. (43)}
$\psi_j^{\mathrm{init}} = c_j - \lambda_j^2$ {cf. eq. (44)}
if $\min_j(\psi_j^{\mathrm{init}}) \leq 0$ **then**
$\psi_j^{\mathrm{init}} = 0.1 \, c_k + c_j - \frac{c_k \, \lambda_j^2}{\lambda_k^2} \; ; \; k = \mathrm{argmin_j} \left(\psi_j^{\mathrm{init}}\right)$
end if
$\sigma_\lambda^2 = \rho \, \mathrm{mean}\,(c_j)$ {cf. eq. (30)}
END initialization

BEGIN factor analysis
while cyc < $\mathrm{cyc}_{\max}$ **do**
{**E-Step** }
FOR j in 1:probes $\phi_j = \frac{\lambda_j}{\psi_j}$ **END FOR** {cf. eq. (35)}
$a = \left(1 + \boldsymbol{\lambda}^T \, \boldsymbol{\phi}\right)^{-1}$ {cf. eq. (36)}
$\boldsymbol{\eta} = \boldsymbol{\phi} \, a$ {cf. eq. (37)}
$\boldsymbol{\zeta} = \boldsymbol{C} \, \boldsymbol{\eta}$ {cf. eq. (38)}
$\xi = a + \boldsymbol{\eta}^T \, \boldsymbol{\zeta}$ {cf. eq. (38)}
{**M-Step**}
FOR j in 1:probes $\lambda_j = \frac{\zeta_j + \psi_j \, \mu_\lambda \, \sigma_\lambda^{-2}}{\xi + \psi_j \, \sigma_\lambda^{-2}}$ **END FOR** {cf. eq. (40) and (28)}
$\boldsymbol{\lambda} = \frac{1}{2}\,(\boldsymbol{\lambda} + |\boldsymbol{\lambda}|)$ {cf. eq. (28)}
FOR j in 1:probes
$\psi_j = c_j - \zeta_j \lambda_j + \psi_j \, \sigma_\lambda^{-2} \, \lambda_j \,(\mu_\lambda - \lambda_j)$ {cf. eq. (28) and (28)}
END FOR
if $(\left\|\boldsymbol{\lambda} - \boldsymbol{\lambda}^{\mathrm{old}}\right\| < \mathrm{tol})$ **then** break while-loop
$\boldsymbol{\lambda}^{\mathrm{old}} = \boldsymbol{\lambda}$ and cyc = cyc + 1
end while
END factor analysis

BEGIN signal extraction
$z_i = \boldsymbol{\eta}^T \boldsymbol{x}_i$ {cf. eq. (33)}
$s_i = \mathrm{mean}\,(\lambda_j) \; z_i \; f + \mathrm{mean}\,(\mu_j)$ {the final signal}
END signal extraction
END FORALL genes

3 Detailed *Affycomp II* Assessment Results

In this appendix we report the detailed results as obtained from *affycomp* **R** package Cope et al. (2004) from the Bioconductor Project (`www.bioconductor.org`). The evaluation is reported in tables which entries are described in the following.

Table Entries. The following list summarizes the table entries, where the first entry is the number of the entry, the optimal possible value is given as "{best:}".

(1) Median of the standard deviation of replicates {best: 0} (data sets A2, B, C).

(2) Average R^2 between replicates {best: 1} (data set A2). R^2 is the correlation coefficient between replicates.

(3) Correlation between fold-change measurements with the lowest 1.25 μg and highest 20 μg RNA concentration {best: 1} (data set A2).

(4) Median of the slope of the curve "observed abundance against the inverse of the dilution" {best: 1} (data set A2).

(5) Averaged slope of the curve "observed concentration against nominal concentration" {best: 1} (data sets A1, B, C).

(6) R^2 of (5) {best: 1} (cf. Plot (d); data sets A1, B, C).

(7) Area under the ROC curve (AUC) for all fold-changes up to 100 false positives {best: 1} (cf. Plot (f-1); data set A1).

(8) False positives, i.e. 1 - specificity {best: 0} (data set A1).

(9) True positives, i.e. sensitivity {best: 16} (data set A1).

(10) Interquartile range of log ratios for non-differentially expressed genes {best: 0} (data sets A1).

(11) Average slope of observed fold-changes are plotted against log true fold-changes {best: 1} (cf. Plot (g-1); data sets A1, B, C).

(12) Like (11), but for low (under 2 pM) RNA concentrations {best: 1} data sets A1, B, C).

(13) Like (7), only for fold-change = 2, i.e. the RNA concentration for the corresponding gene was doubled which is the minimal increase {best: 1} (data sets A1).

(14) Like (8), but for fold-change = 2 {best: 0} (data set A1).

(15) Like (9), but for fold-change = 2 {best: 16} (data set A1).

(16) Interquartile range of the log fold-changes from non-differentially expressed genes {best: 0} (data sets B, C).

(17) 99.9% percentile of the log fold-changes from non-differentially expressed genes {best: 0} (data sets B, C).

(18) Slope of regression function for observed versus nominal log concentration for genes with low intensities {best: 1} (data sets B, C).

(19) Like (18), but for genes with medium intensities {best: 1} (data sets B, C).

(20) Like (19), but for genes with high intensities {best: 1} (data sets B, C).

(21) Area under the ROC (AUC) curve (up to 100 false positives) for genes with low intensity standardized so that optimum is 1 {best: 1} (data sets B, C).

(22) Like (21), but for genes with medium intensities {best: 1} (data sets B, C).

(23) Like (22), but for genes with high intensities {best: 1} (data sets B, C).

(24) Weighted average of the AUC values (21), (22), and (23) with weights related to amount of data in each class {best: 1} (data sets B, C).

Table 1 (data set A), Table 2 (data set B), and Table 3 (data set C) summarize the full results for data sets A, B, and C.

N	Method-Submitter	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	(perfection)	*0.00*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*	*0.00*	*16.00*	*0.00*	*1.00*	*1.00*	*1.00*	*0.00*	*16.00*
1	MAS_5.0-rafa	0.29	0.89	0.73	0.85	0.71	0.86	0.36	3108.99	12.82	2.66	0.69	0.65	0.07	3072.18	3.71
2	RMA-rafa	0.09	0.99	0.94	0.87	0.63	0.80	0.82	15.84	11.98	0.31	0.61	0.36	0.54	1.00	1.71
3	dChip-rafa	0.09	0.99	0.91	0.77	0.53	0.85	0.67	36.91	11.43	0.45	0.52	0.32	0.17	28.64	1.25
4	ZAM2NBG-astrand	0.07	0.99	0.94	0.72	0.57	0.77	0.84	2.44	11.70	0.24	0.57	0.32	0.61	0.57	1.14
5	qn.p5-cope	0.11	0.98	0.56	0.06	0.42	0.50	0.62	20.30	9.58	0.38	0.43	0.14	0.24	15.75	1.39
6	vsn_scal-w.huber	0.08	0.99	0.96	**1.00**	0.77	0.81	0.85	6.69	12.23	0.23	0.75	0.28	0.66	0.43	3.89
7	vsn-w.huber	0.06	0.99	0.96	0.67	0.51	0.81	0.85	0.40	10.83	0.15	0.50	0.19	0.66	0.21	1.11
8	RMAVSN-cappola	0.09	0.99	0.94	0.89	0.61	0.81	0.83	17.87	11.79	0.25	0.60	0.32	0.59	0.50	1.61
9	RMA_NBG-bolstad	0.04	**1.00**	0.91	0.56	0.48	0.81	0.85	0.13	10.45	0.12	0.47	0.15	0.68	**0.11**	1.04
10	GSVDmin-hzuzan	0.05	0.98	0.97	0.59	0.50	0.83	0.81	4.87	11.09	0.21	0.49	0.24	0.56	2.43	1.00
11	PLIER-Hubbell	0.13	0.09	0.01	0.84	0.71	0.91	0.02	596.77	12.85	4.03	0.72	0.65	0.02	589.96	3.57
12	GSVDmod-hzuzan	0.05	**1.00**	0.97	0.55	0.51	0.85	0.84	0.79	11.19	0.19	0.50	0.24	0.60	0.54	1.11
13	PLIER+16-Hubbell	0.08	0.99	0.88	0.64	0.65	**0.91**	0.81	8.42	12.34	0.34	0.65	0.46	0.46	5.07	2.04
14	GCRMA-zwu	0.09	0.99	0.89	0.72	0.97	0.84	0.82	7.62	12.97	0.35	0.92	0.66	0.54	7.07	5.29
15	ChipMan-plauren	0.31	0.99	0.94	1.26	0.88	0.82	0.67	183.99	13.03	0.67	0.87	0.44	0.20	159.86	5.11
16	ProbePro-shilmer	0.16	0.70	0.58	0.84	1.45	0.47	0.17	2087.07	12.53	15.70	1.33	1.93	0.07	2046.46	4.93
17	MMEI-shibing.deng	0.02	**1.00**	0.92	0.52	0.45	0.80	0.86	0.12	10.41	0.12	0.45	0.16	0.69	**0.11**	1.00
18	PM-zhangli	0.05	0.99	0.97	0.53	0.46	0.87	0.84	1.39	10.67	0.15	0.45	0.18	0.64	0.68	1.00
19	RMA:GNV-szeto	0.09	0.99	0.98	0.68	0.62	0.80	0.82	15.86	11.99	0.31	0.61	0.36	0.54	1.00	1.71
20	GL-mai98ftu	0.05	0.99	0.92	0.56	0.48	0.81	0.83	0.15	10.42	0.14	0.47	0.16	0.66	**0.11**	1.18
21	MAS5+32-Hubbell	0.07	0.98	0.93	0.71	0.60	0.88	0.72	20.56	11.76	0.51	0.59	0.33	0.18	19.18	1.68
22	gMOS_v.1-m.milo	0.32	0.97	0.81	0.64	0.95	0.75	0.54	1358.01	12.75	2.15	0.94	**1.04**	0.10	1319.07	5.36
23	rsvd-jack liu	–	–	–	–	0.66	0.90	0.81	2.63	12.14	0.35	0.66	0.41	0.49	2.04	2.89
24	ZL-lzhou	0.34	0.99	0.04	0.23	0.57	0.65	0.79	21.99	11.98	0.24	0.57	0.36	0.42	37.57	2.36
32	gltran-lzhou	0.04	0.99	0.94	0.64	0.51	0.78	0.84	1.46	11.09	0.19	0.50	0.22	0.65	0.68	1.04
33	UM-Tr-Mn-jmacdon	–	–	–	–	0.68	0.87	0.51	1399.72	12.56	1.76	0.67	0.54	0.07	1385.00	2.64
34	GS_RMA-thon	–	–	–	–	0.63	0.80	0.82	15.86	11.98	0.31	0.61	0.36	0.54	0.93	1.71
35	GS_GCRMA-thon	–	–	–	–	0.84	**0.91**	0.84	6.53	13.15	0.41	0.82	0.65	0.58	3.00	4.71
36	gcrma113-zwu	–	–	–	–	0.87	**0.91**	0.85	3.89	13.16	0.37	0.86	0.68	0.61	2.57	4.89
41	mgMOS_gs-liux	0.21	0.96	0.88	0.80	0.76	0.82	0.57	1061.31	13.12	1.37	0.75	0.90	0.07	1028.61	4.14
42	mmgMOSgs-liux	0.23	0.96	0.82	0.86	**1.03**	0.80	0.59	1616.01	**13.65**	1.79	**1.02**	1.40	0.09	1570.11	**6.71**
43	vsn-huber	0.06	0.99	0.96	0.67	0.51	0.81	0.84	0.38	10.81	0.15	0.50	0.18	0.66	0.21	1.11
44	vsn_scal-huber	0.08	0.99	0.96	1.00	0.76	0.81	0.84	6.45	12.20	0.23	0.74	0.28	0.66	0.39	3.82
49	dChipwMM-rafa	0.42	0.86	0.73	0.83	0.91	0.61	0.12	872.82	12.81	3.11	0.87	1.68	0.07	842.21	4.18
50	NLF95-scvortso +	–	–	–	–	0.65	0.89	**0.91**	4.90	11.89	**0.00**	0.64	0.34	**0.91**	0.82	1.25
53	Afit_qn-Michael.D.Lynch +	–	–	–	–	0.79	0.82	0.76	61.60	12.84	0.57	0.75	0.66	0.34	31.82	3.68
54	Afit_fn-Michael.D.Lynch +	–	–	–	–	0.79	0.80	0.72	767.26	12.94	0.55	0.75	0.68	0.23	437.00	4.29
56	PQN-wei-min.liu +	–	–	–	–	0.57	0.83	0.78	17.68	11.64	0.35	0.56	0.33	0.36	3.86	1.07
57	PQNB1-wei-min.liu +	–	–	–	–	0.08	0.78	0.75	**0.00**	0.21	0.05	0.08	0.07	0.33	**0.00**	0.29
58	PQNB1+16-wei-min.liu +	–	–	–	–	0.01	0.81	0.79	**0.00**	0.00	**0.00**	0.01	0.01	0.41	**0.00**	0.00
59	DQN-wei-min.liu +	–	–	–	–	0.72	0.88	0.56	1693.28	12.82	1.48	0.71	0.68	0.07	1666.57	3.46
60	DQNB1-wei-min.liu +	–	–	–	–	0.08	0.84	0.56	**0.00**	0.44	0.15	0.07	0.10	0.08	**0.00**	0.64
61	DQNB1+16-wei-min.liu +	–	–	–	–	0.01	0.87	0.64	**0.00**	0.00	0.01	0.01	0.01	0.09	**0.00**	0.00
62	l-FARMS-hochreit	**0.00**	**1.00**	0.96	1.29	0.77	0.80	0.85	0.34	11.82	0.03	0.80	0.31	0.78	0.25	5.04
63	q-FARMS-hochreit	**0.00**	**1.00**	**0.99**	0.81	0.59	0.78	0.89	0.17	10.87	**0.00**	0.60	0.23	0.84	0.11	2.32
0	(perfection)	*0.00*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*	*0.00*	*16.00*	*0.00*	*1.00*	*1.00*	*1.00*	*0.00*	*16.00*

Table 1: Result table for data set A taken from affycomp website. The columns are described above; the best result is in bold.

N	Method-Submitter	1	16	17	6	5	18	19	20	11	12	21	22	23	24
0	(perfection)	*0.00*	*0.00*	*0.00*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*
1	MAS_5.0-rafa	0.63	0.85	4.48	0.86	0.71	0.72	0.80	0.45	0.69	0.65	0.07	0.00	0.00	0.05
2	RMA-rafa	0.11	0.19	0.57	0.80	0.63	0.29	0.73	0.47	0.61	0.36	0.51	0.91	0.64	0.60
3	dChip-rafa	0.13	0.20	1.44	0.85	0.53	0.25	0.64	0.39	0.52	0.32	0.21	0.43	0.16	0.26
4	ZAM2NBG-magnus.astrand	0.09	0.16	0.50	0.77	0.57	0.25	0.66	0.47	0.57	0.32	0.57	0.95	0.74	0.66
5	qn.p5-cope	0.12	0.22	1.09	0.50	0.42	0.11	0.44	0.52	0.43	0.14	0.09	0.43	0.46	0.17
6	vsn_scal-w.huber	0.09	0.15	0.43	0.81	0.77	0.21	0.82	0.70	0.75	0.28	0.53	0.97	0.86	0.64
7	vsn-w.huber	0.06	0.10	0.29	0.81	0.51	0.14	0.55	0.47	0.50	0.19	0.53	0.97	0.86	0.64
8	RMAVSN-thomas.cappola	0.09	0.16	0.48	0.81	0.61	0.25	0.71	0.46	0.60	0.32	0.52	0.94	0.70	0.62
9	RMA_NBG-bolstad	0.04	0.08	0.24	0.81	0.48	0.12	0.50	0.46	0.47	0.15	0.55	0.97	0.92	0.65
10	GSVDmin-hzuzan	0.08	0.13	0.60	0.83	0.50	0.17	0.58	0.41	0.49	0.24	0.39	0.87	0.65	0.51
11	PLIER-Earl Hubbell	0.19	0.33	123.27	**0.91**	0.71	0.76	0.84	0.46	0.72	0.65	0.04	0.00	0.00	0.03
12	GSVDmod-hzuzan	0.07	0.13	0.44	0.85	0.51	0.18	0.60	0.42	0.50	0.24	0.47	0.94	0.74	0.59
13	PLIER+16-Earl Hubbell	0.13	0.21	0.83	**0.91**	0.65	0.45	0.78	0.46	0.65	0.46	0.61	0.83	0.46	0.66
14	GCRMA-zwu	0.09	0.16	0.77	0.84	0.97	0.73	1.19	0.55	0.92	0.66	0.62	0.94	0.59	0.69
15	ChipMan-plauren	0.27	0.33	2.26	0.82	0.88	0.36	1.04	0.68	0.87	0.44	0.21	0.50	0.24	0.28
16	ProbePro-shilmer	0.31	0.47	18.75	0.47	1.45	1.26	1.73	0.39	1.33	1.93	0.05	0.04	0.00	0.05
17	MMEI-shibing.deng	0.04	0.06	0.23	0.80	0.45	0.11	0.49	0.46	0.45	0.16	0.57	0.98	0.94	0.67
18	PM-zhangli *	0.05	0.09	0.40	0.87	0.46	0.14	0.52	0.43	0.45	0.18	0.47	0.93	0.83	0.59
19	RMA:GNV-szeto	0.11	0.19	0.58	0.80	0.62	0.29	0.73	0.47	0.61	0.36	0.50	0.91	0.64	0.60
20	GL-mai98ftu	0.05	0.08	0.25	0.81	0.48	0.12	0.50	0.46	0.47	0.16	0.58	0.96	0.88	0.67
21	MAS5+32-Earl Hubbell	0.14	0.23	1.07	0.88	0.60	0.31	0.68	0.44	0.59	0.33	0.04	0.28	0.08	0.10
22	gMOS_v.1-m.milo	0.29	**0.00**	3.35	0.75	0.95	0.82	1.22	0.42	0.94	**1.04**	0.07	0.04	0.00	0.06
23	rsvd-jack liu	**0.00**	**0.00**	0.58	0.90	0.66	0.31	0.84	0.40	0.66	0.41	0.54	0.93	0.54	0.64
24	ZL-lzhou	0.22	0.12	0.52	0.65	0.57	0.39	0.67	0.45	0.57	0.36	0.65	0.87	0.73	0.70
32	gltran-lzhou	0.07	0.12	0.42	0.78	0.51	0.19	0.57	0.45	0.50	0.22	0.55	0.94	0.80	0.65
33	UM-Tr-Mn-jmacdon	0.32	0.51	2.92	0.87	0.68	0.53	0.82	0.42	0.67	0.54	0.11	0.00	0.00	0.08
34	GS_RMA-thon	0.11	0.19	0.57	0.80	0.63	0.29	0.73	0.47	0.61	0.36	0.51	0.91	0.64	0.60
35	GS_GCRMA-thon	0.10	0.07	0.79	**0.91**	0.84	0.51	**1.02**	0.55	0.82	0.65	0.64	0.94	0.56	0.72
36	gcrma113-zwu	0.08	0.04	0.74	**0.91**	0.87	0.52	1.06	0.56	0.86	0.68	0.68	0.97	0.63	0.75
41	mgMOS_gs-liux	0.36	0.55	2.86	0.82	0.76	0.77	0.89	0.43	0.75	0.90	0.25	0.04	0.00	0.20
42	mmgMOSgs-liux	0.40	0.58	3.27	0.80	**1.03**	**1.21**	1.26	0.45	**1.02**	1.40	0.36	0.07	0.00	0.29
43	vsn-huber	0.06	0.10	0.28	0.81	0.51	0.14	0.54	0.46	0.50	0.18	0.53	0.97	0.86	0.64
44	vsn_scal-huber	0.09	0.15	0.42	0.81	0.76	0.21	0.82	0.70	0.74	0.28	0.53	0.97	0.86	0.64
49	dChipwMM-rafa	0.32	0.38	10.83	0.61	0.91	0.78	0.97	0.34	0.87	1.68	0.10	0.00	0.00	0.08
50	NLF95-scvortso	**0.00**	**0.00**	**0.00**	0.89	0.65	0.25	0.68	0.68	0.64	0.34	0.74	0.92	0.95	0.79
53	Afit_qn-Michael.D.Lynch	0.07	0.00	1.18	0.82	0.79	0.62	0.88	0.50	0.75	0.66	0.62	0.70	0.32	0.64
54	Afit_fn-Michael.D.Lynch	0.39	0.15	1.20	0.80	0.79	0.65	0.91	0.49	0.75	0.68	0.59	0.64	0.33	0.60
56	PQN-wei-min.liu	0.11	0.19	0.75	0.83	0.57	0.26	0.66	0.45	0.56	0.33	0.39	0.80	0.47	0.49
57	PQNB1-wei-min.liu	0.02	0.03	0.11	0.78	0.08	0.05	0.11	0.04	0.08	0.07	0.50	0.71	0.14	0.55
58	PQNB1+16-wei-min.liu	**0.00**	**0.00**	0.01	0.81	0.01	0.00	0.01	0.00	0.01	0.01	0.48	0.83	0.37	0.56
59	DQN-wei-min.liu	0.47	0.68	2.59	0.88	0.72	0.68	0.87	0.41	0.71	0.68	0.23	0.01	0.00	0.18
60	DQNB1-wei-min.liu	0.05	0.08	0.27	0.84	0.08	0.09	0.09	0.04	0.07	0.10	0.35	0.06	0.00	0.28
61	DQNB1+16-wei-min.liu	**0.00**	0.01	0.02	0.87	0.01	0.01	0.01	0.00	0.01	0.01	0.38	0.17	0.01	0.33
62	l-FARMS-hochreit	**0.00**	**0.00**	0.19	0.80	0.77	0.21	0.90	**0.78**	0.80	0.31	0.80	0.95	0.94	0.84
63	q-FARMS-hochreit	**0.00**	**0.00**	0.10	0.78	0.59	0.15	0.68	0.59	0.60	0.23	**0.89**	0.97	**0.97**	**0.91**
0	(perfection)	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table 2: Result table for data set B taken from affycomp website. The columns are described above; the best result is in bold.

N	Method-Submitter	1	16	17	6	5	18	19	20	11	12	21	22	23	24
0	(perfection)	*0.00*	*0.00*	*0.00*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*
1	MAS_5.0-rafa	0.29	0.47	4.01	0.91	0.77	0.58	0.73	0.77	0.77	0.64	0.09	0.00	0.00	0.06
2	RMA-rafa	0.07	0.13	0.40	0.90	0.68	0.20	0.71	0.80	0.68	0.31	0.57	0.91	0.96	0.65
8	RMAVSN-thomas.cappola	0.02	0.04	0.15	0.89	0.12	0.06	0.13	0.10	0.12	0.08	0.46	0.59	0.43	0.49
23	rsvd-jack liu	0.14	0.12	0.73	**0.94**	0.74	0.31	0.78	0.73	0.74	0.43	0.53	0.73	0.71	0.58
25	rsvd.pm-jack liu	0.06	0.11	0.34	0.89	0.53	0.12	0.53	0.77	0.53	0.16	0.42	0.90	0.96	0.54
26	rma-tog-dgreco	0.07	0.13	0.40	0.90	0.68	0.20	0.71	0.80	0.68	0.31	0.57	0.91	0.96	0.65
27	rma-sep-dgreco	0.18	0.28	0.96	0.90	0.71	0.27	0.72	0.84	0.71	0.39	0.38	0.53	0.63	0.42
28	LW1-dgreco	0.08	0.14	1.18	0.91	0.59	0.19	0.62	0.74	0.59	0.25	0.23	0.47	0.55	0.29
29	LW2-dgreco	0.14	0.25	13.88	0.56	1.08	1.50	0.80	0.68	1.08	1.45	0.19	0.00	0.00	0.14
30	rsvd.bgc-jack liu	0.08	0.14	0.52	0.89	0.58	0.16	0.59	0.79	0.58	0.22	0.38	0.80	0.90	0.49
31	cor523-cope	0.02	0.03	0.12	0.88	0.12	0.06	0.13	0.10	0.12	0.08	0.54	0.77	0.61	0.60
33	UM-Tr-Mn-jmacdon	0.15	0.25	1.86	0.93	0.70	0.36	0.72	0.70	0.70	0.44	0.18	0.10	0.10	0.16
34	GS_RMA-thon	0.07	0.13	0.40	0.90	0.68	0.20	0.71	0.80	0.68	0.30	0.56	0.91	0.96	0.65
35	GS_GCRMA-thon	0.07	0.09	0.65	0.93	0.93	0.37	0.96	0.96	0.93	0.55	0.59	0.87	0.90	0.66
36	gcrma113-zwu	0.06	0.04	0.61	0.91	**1.00**	0.25	1.13	**0.97**	**1.00**	0.48	0.45	0.91	0.92	0.57
37	rsvd2-jack.liu	0.17	0.28	1.74	0.91	0.75	0.46	0.74	0.81	0.75	0.52	0.29	0.16	0.21	0.26
38	W237-dario.greco	0.02	0.04	0.17	0.87	0.12	0.05	0.13	0.10	0.12	0.07	0.35	0.54	0.39	0.39
39	RMA_NBG-bolstad	0.01	0.02	0.06	0.90	0.09	0.02	0.09	0.10	0.09	0.04	0.54	0.90	0.93	0.63
40	RMAVSN-thomas.cappola	0.02	0.04	0.15	0.89	0.12	0.06	0.13	0.10	0.12	0.08	0.46	0.59	0.43	0.49
41	mgMOS_gs-liux	0.24	0.34	2.02	0.92	0.81	0.58	0.78	0.77	0.81	0.68	0.33	0.08	0.12	0.27
42	mmgMOSgs-liux	0.24	0.34	2.71	0.89	1.03	**0.88**	**0.98**	0.79	1.03	**1.05**	0.37	0.02	0.01	0.28
43	vsn-huber	0.05	0.09	0.24	0.88	0.52	0.10	0.51	0.78	0.52	0.16	0.53	0.93	0.99	0.63
44	vsn_scal-huber	0.07	0.13	0.36	0.88	0.78	0.15	0.77	1.16	0.78	0.23	0.53	0.93	0.99	0.63
45	NLF-scvortso	0.12	0.18	1.79	0.86	0.70	0.21	0.67	1.10	0.70	0.28	0.20	0.31	0.59	0.23
46	real-scvortso	0.05	0.02	0.51	0.00	-0.01	-0.06	-0.02	0.06	-0.01	-0.00	0.55	0.40	0.42	0.51
47	real-scvortso	0.05	0.02	0.53	0.90	1.04	0.09	1.34	**0.97**	1.04	0.30	0.33	0.95	0.93	0.48
48	NLF-scvortso	0.02	0.01	0.15	0.85	0.70	0.22	0.69	1.07	0.70	0.29	0.76	0.95	0.97	0.81
49	dChipwMM-rafa	0.20	0.26	9.61	0.65	0.98	1.13	0.81	0.69	0.99	1.17	0.19	0.00	0.00	0.14
51	Afit_qn-Michael.D.Lynch	0.06	0.03	0.74	0.87	0.86	0.55	0.86	0.82	0.86	0.66	0.61	0.78	0.79	0.65
52	Afit_fn-Michael.D.Lynch	0.18	0.17	0.79	0.87	0.86	0.55	0.85	0.82	0.86	0.67	0.61	0.74	0.73	0.64
55	SBLW-iliketobicycle	0.09	0.15	0.68	0.90	0.62	0.23	0.65	0.78	0.62	0.28	0.44	0.70	0.79	0.51
56	PQN-wei-min.liu	0.07	0.14	0.64	0.91	0.59	0.18	0.60	0.78	0.59	0.24	0.35	0.73	0.85	0.44
57	PQNB1-wei-min.liu	0.01	0.02	0.09	0.92	0.09	0.03	0.11	0.10	0.09	0.04	0.42	0.81	0.83	0.51
58	PQNB1+16-wei-min.liu	**0.00**	**0.00**	**0.01**	0.91	0.01	0.00	0.01	0.01	0.01	0.00	0.40	0.87	0.96	0.51
59	DQN-wei-min.liu	0.24	0.35	2.21	0.93	0.76	0.48	0.76	0.72	0.76	0.57	0.22	0.03	0.05	0.17
60	DQNB1-wei-min.liu	0.02	0.03	0.15	**0.94**	0.09	0.04	0.09	0.10	0.09	0.05	0.32	0.43	0.48	0.34
61	DQNB1+16-wei-min.liu	**0.00**	**0.00**	**0.01**	0.92	0.01	0.00	0.01	0.01	0.01	0.00	0.33	0.61	0.83	0.40
62	l-FARMS-hochreit	**0.00**	**0.00**	0.08	0.90	1.02	0.23	1.04	1.49	1.02	0.33	0.91	**0.99**	**1.00**	0.93
63	q-FARMS-hochreit	**0.00**	**0.00**	0.04	0.89	0.77	0.17	0.78	1.12	0.77	0.25	**0.94**	**0.99**	**1.00**	**0.95**
0	(perfection)	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table 3: Result table for data set C taken from affycomp website. The columns are described above; the best result is in bold.

References

L. M. Cope, R. A. Irizarry, H. A. Jaffee, Z. Wu, and T. P. Speed. A benchmark for Affymetrix GeneChip expression measures. *Bioinformatics*, 20(3):323–331, 2004.

M. H. DeGroot. *Optimal Statistical Decisions*. McGraw-Hill, New York, 1970.

A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society B*, 39(1):1–22, 1977.

B. S. Everitt. *An introduction to latent variable models*. Chapman and Hall, London, 1984.

J.-L. Gauvain and C.-H. Lee. Maximum a posteriori estimation for multivariate gaussian mixture observations of Markov chains. *IEEE Transactions on Speech and Audio Processing*, 2(2): 291–298, 1994.

K. G. Jöreskog. Some contributions to maximum likelihood factor analysis. *Psychometrika*, 32: 443–482, 1967.

D. Rubin and D. Thayer. EM algorithms for ML factor analysis. *Psychometrika*, 47(1):69–76, 1982.