A Fully Bayesian Model for Gene Expression Heterosis in RNA-seq Data

Will Landau

Department of Statistics
Iowa State University

February 23, 2014

Contents

1 Introduction

This writeup explains a fully Bayesian Markov chain Monte Carlo method for modeling RNA-seq data. The hierarchical model featured focuses on heterosis, or hybrid vigor, a phenomenon that concerns two parental genetic lines and an hybrid line. For each gene in an RNA-seq dataset, we consider three types of heterosis at the level of gene expression:

1. High parent heterosis: the gene is significantly more expressed in the hybrid than in either of the parent lines.
2. Low parent heterosis: the gene is significantly less expressed in the hybrid than in either of the parent lines.
3. Mid parent heterosis: the expression level of the gene in the hybrid is significantly different from the average of the parental expression levels.

Let $y_{g,n}$ be the expression level of gene g ($g = 1, \ldots, G$) in sample n ($n = 1, \ldots, N$). The samples come from one of three groups: group 1, the first parent, group 2, the hybrid, and group 3, the second parent. Hence, we define:

- μ_{g1}: mean expression level of gene g in the first parent
- μ_{g2}: mean expression level of gene g in the hybrid
- μ_{g3}: mean expression level of gene g in the second parent

In the model below, there are three quantities of primary interest:

- $\phi_g = \frac{\mu_{g1}+\mu_{g3}}{2}$, the parental mean expression level of gene g.
- $\alpha_g = \frac{\mu_{g1}-\mu_{g3}}{2}$, half the parental difference in expression levels of gene g.
- $\delta_g = \mu_{g2} - \phi_g$, the overexpression of gene g in the hybrid relative to the parental mean.

With MCMC samples of these quantities, for some threshold $\varepsilon > 0$, we can calculate empirical estimates of the following probabilities of interest:

- $P(|\alpha_g| \geq \varepsilon \mid \boldsymbol{y})$, the probability of differential expression.
- $P(\delta_g > |\alpha_g| \mid \boldsymbol{y})$, the probability of high parent heterosis.
- $P(\delta_g < -|\alpha_g| \mid \boldsymbol{y})$, the probability of low parent heterosis.
- $P(|\delta_g| \geq \varepsilon \mid \boldsymbol{y})$, the probability of mid parent heterosis.

2 The Model

$$
\begin{aligned}
y_{g,n} &\overset{\text{ind}}{\sim} \text{Poisson}(\exp(\rho_n + \varepsilon_{g,n} + \eta(g,n))) \\
\rho_n &\overset{\text{ind}}{\sim} \text{N}(0, \sigma_\rho^2) \\
\sigma_\rho &\sim \text{U}(0, s_\rho) \\
\varepsilon_{g,n} &\overset{\text{ind}}{\sim} \text{N}(0, \gamma_g^2) \\
\gamma_g^2 &\overset{\text{ind}}{\sim} \text{Inv-Gamma}\left(\text{shape} = \frac{\nu}{2}, \ \text{scale} = \frac{\nu \cdot \tau^2}{2}\right) \\
\nu &\sim \text{U}(0, d) \\
\tau^2 &\sim \text{Gamma}(\text{shape} = a, \text{rate} = b) \\
\phi_g &\overset{\text{ind}}{\sim} \text{N}(\theta_\phi, \sigma_\phi^2) \\
\theta_\phi &\sim \text{N}(0, c_\phi^2) \\
\sigma_\phi &\sim \text{U}(0, s_\phi) \\
\alpha_g &\overset{\text{ind}}{\sim} \text{N}(\theta_\alpha, \sigma_\alpha^2) \\
\theta_\alpha &\sim \text{N}(0, c_\alpha^2) \\
\sigma_\alpha &\sim \text{U}(0, s_\alpha) \\
\delta_g &\overset{\text{ind}}{\sim} \text{N}(\theta_\delta, \sigma_\delta^2) \\
\theta_\delta &\sim \text{N}(0, c_\delta^2) \\
\sigma_\delta &\sim \text{U}(0, s_\delta)
\end{aligned}
$$

where:

- Conditional independence is implied unless otherwise specified.
- The parameters to the left of the "$\sim$" are implicitly conditioned on the parameters to the right.
- $\eta(g, n)$ is the function given by:

$$
\eta(g,n) = \begin{cases} \phi_g - \alpha_g & \text{library } n \text{ is in treatment group 1 (parent 1)} \\ \phi_g + \delta_g & \text{library } n \text{ is in treatment group 2 (hybrid)} \\ \phi_g + \alpha_g & \text{library } n \text{ is in treatment group 3 (parent 2)} \end{cases}
$$

$$
\eta(g,n) = \begin{cases} \phi_g - \alpha_g & \text{sample } n \text{ from parent 1 genotype} \\ \phi_g + \delta_g & \text{sample } n \text{ from hybrid genotype} \\ \phi_g + \alpha_g & \text{sample } n \text{ from parent 2 genotype} \end{cases}
$$

3 Full Conditional Distributions

Define:

- $k(n) =$ treatment group of library n.
- $\lambda_{g,n} = \exp(\rho_n + \varepsilon_{g,n} + \eta(g,n))\ )$
- $G_\alpha =$ number of genes for which $\alpha_g \neq 0$
- $G_\delta =$ number of genes for which $\delta_g \neq 0$

Then:

$$p(\nu \mid \cdots) \propto \Gamma\left(\nu/2\right)^{-G} \left(\frac{\nu \cdot \tau^2}{2}\right)^{G\nu/2} \left(\prod_{g=1}^{G} \gamma_g^2\right)^{-(\nu/2+1)} \exp\left(-\frac{\nu \cdot \tau^2}{2} \sum_{g=1}^{G} \frac{1}{\gamma_g^2}\right) I(0 < \nu < d)$$

$$p(\rho_n \mid \cdots) \propto \exp\left(\rho_n G \overline{y}_{.n} - \frac{\rho_n^2}{2\sigma_\rho^2} - \exp(\rho_n) \sum_{g=1}^{G} \exp(\varepsilon_{g,n} + \eta(g,n))\right)$$

$$p(\phi_g \mid \cdots) \propto \exp\left(\phi_g N \overline{y}_{g.} - \frac{(\phi_g - \theta_\phi)^2}{2\sigma_\phi^2} - \exp(\phi_g)\left[\exp(-\alpha_g) \sum_{k(n)=1} \exp(\rho_n + \varepsilon_{g,n}) \right.\right.$$
$$\left.\left. + \exp(\delta_g) \sum_{k(n)=2} \exp(\rho_n + \varepsilon_{g,n}) + \exp(\alpha_g) \sum_{k(n)=3} \exp(\rho_n + \varepsilon_{g,n})\right]\right)$$

$$p(\alpha_g \mid \cdots) \propto \exp\left(\alpha_g \left(\sum_{k(n)=3} y_{g,n} - \sum_{k(n)=1} y_{g,n}\right) - \frac{(\alpha_g - \theta_\alpha)^2}{2\sigma_\alpha^2}\right.$$
$$- \exp(\alpha_g)\exp(\phi_g) \sum_{k(n)=3} \exp(\rho_n + \varepsilon_{g,n})$$
$$\left. - \exp(-\alpha_g)\exp(\phi_g) \sum_{k(n)=1} \exp(\rho_n + \varepsilon_{g,n})\right)$$

$$p(\delta_g \mid \cdots) \propto \exp\left(\delta_g \sum_{k(n)=2} y_{g,n} - \frac{(\delta_g - \theta_\delta)^2}{2\sigma_\delta^2} - \exp(\delta_g)\exp(\phi_g) \sum_{k(n)\neq 2} \exp(\rho_n + \varepsilon_{g,n})\right)$$

$$p(\varepsilon_{g,n} \mid \cdots) \propto \exp\left(y_{g,n}\varepsilon_{g,n} - \frac{\varepsilon_{g,n}^2}{2\gamma_g^2} - \exp(\varepsilon_{g,n})\exp(\rho_n + \eta(g,n))\right)$$

$$p(\theta_\phi \mid \cdots) = \mathrm{N}\left(\theta_\phi \;\middle|\; \frac{c_\phi^2 \sum_{g=1}^G \phi_g}{Gc_\phi^2 + \sigma_\phi^2},\ \frac{c_\phi^2 \sigma_\phi^2}{Gc_\phi^2 + \sigma_\phi^2}\right)$$

$$p(\theta_\alpha \mid \cdots) = N\left(\theta_\alpha \;\middle|\; \frac{c_\alpha^2 \sum_{g=1}^G \alpha_g}{G_\alpha c_\alpha^2 + \sigma_\alpha^2},\ \frac{c_\alpha^2 \sigma_\alpha^2}{G_\alpha c_\alpha^2 + \sigma_\alpha^2}\right)$$

$$p(\theta_\delta \mid \cdots) = N\left(\theta_\delta \;\middle|\; \frac{c_\delta^2 \sum_{g=1}^G \delta_g}{G_\delta c_\delta^2 + \sigma_\delta^2},\ \frac{c_\delta^2 \sigma_\delta^2}{G_\delta c_\delta^2 + \sigma_\delta^2}\right)$$

$$p\left(\frac{1}{\sigma_\phi^2} \;\middle|\; \cdots\right) = \mathrm{Gamma}\left(\text{shape} = \frac{G-1}{2},\ \text{rate} = \frac{1}{2}\sum_{g=1}^G (\phi_g - \theta_\phi)^2\right) \mathrm{I}\left(\frac{1}{\sigma_\phi^2} > \frac{1}{s_\phi^2}\right)$$

$$p\left(\frac{1}{\sigma_\alpha^2} \;\middle|\; \cdots\right) = \mathrm{Gamma}\left(\text{shape} = \frac{G-1}{2},\ \text{rate} = \frac{1}{2}\sum_{g=1}^G (\alpha_g - \theta_\alpha)^2\right) \mathrm{I}\left(\frac{1}{\sigma_\alpha^2} > \frac{1}{s_\alpha^2}\right)$$

$$p\left(\frac{1}{\sigma_\delta^2} \;\middle|\; \cdots\right) = \mathrm{Gamma}\left(\text{shape} = \frac{G-1}{2},\ \text{rate} = \frac{1}{2}\sum_{g=1}^G (\delta_g - \theta_\delta)^2\right) \mathrm{I}\left(\frac{1}{\sigma_\delta^2} > \frac{1}{s_\delta^2}\right)$$

$$p\left(\frac{1}{\sigma_\rho^2} \mid \cdots\right) = \mathrm{Gamma}\left(\frac{1}{\sigma_\rho^2} \mid \text{shape} = \frac{N-1}{2},\ \text{rate} = \frac{1}{2}\sum_{n=1}^N \rho_n^2\right) I\left(\frac{1}{\sigma_\rho^2} > \frac{1}{s_\rho^2}\right)$$

$$p(\tau^2 \mid \cdots) = \mathrm{Gamma}\left(\tau^2 \mid \text{shape} = a + \frac{G\nu}{2},\ \text{rate} = b + \frac{\nu}{2}\sum_{g=1}^G \frac{1}{\gamma_g^2}\right)$$

$$p\left(\frac{1}{\gamma_g^2} \;\middle|\; \cdots\right) = \mathrm{Gamma}\left(\frac{1}{\gamma_g^2} \;\middle|\; \text{shape} = \frac{N+\nu}{2},\ \text{rate} = \frac{1}{2}\left(\nu \cdot \tau^2 + \sum_{n=1}^N \varepsilon_{g,n}^2\right)\right)$$

4 The Gibbs Sampler

Using conditional independence, I can construct Gibbs steps within which I sample parameters simultaneously:

- σ_ρ
- ν
- τ^2
- θ_ϕ
- θ_α
- θ_δ
- σ_ϕ
- σ_α

- σ_δ
- $\gamma_1^2, \ldots, \gamma_g^2$
- $\phi_1, \ldots, \phi_g$
- $\alpha_1, \ldots, \alpha_g$
- $\delta_1, \ldots, \delta_g$
- $\rho_1, \ldots, \rho_n$

Steps with multiple sampled parameters will sample those parameters in parallel on the GPU.

5 Diagnostics

5.1 Gelman Factors

The potential scale reduction factor introduced in the textbook by Gelman **?** monitors the lack of convergence of a single variable in an MCMC. Let η_{ij} be the $i'th$ MCMC draw of a single variable in chain j. Then, the potential scale reduction factor, $\widehat{R}$, compares the within-chain variance, W, to the between-chain variance, B. Suppose there are J chains, each with I iterations. Then,

$$\widehat{R} = \sqrt{1 - \frac{1}{I}\left(\frac{B}{W} - 1\right)}$$

$$B = \frac{I}{J-1}\sum_{j=1}^{J}(\overline{\eta}_{.j} - \overline{\eta}_{..})^2, \qquad \overline{\eta}_{.j} = \frac{1}{I}\sum_{i=1}^{I}\eta_{ij}, \quad \overline{\eta}_{..}\sum_{j=1}^{J}\overline{\eta}_{.j}$$

$$W = \frac{1}{J}\sum_{j=1}^{J}s_j^2, \qquad s_j^2 = \frac{1}{I-1}\sum_{i=1}^{I}(\eta_{ij} - \overline{\eta}_{.j})^2$$

$\widehat{R} \to 1$ as $I \to \infty$. An $\widehat{R}$ value far above 1 indicates a lack of convergence, but an $\widehat{R}$ value near 1 does not imply convergence.

The Gelman factor used in this analysis is not actually the one given above, but a degrees-of-freedom-adjusted version implemented in the `gleman.diag()` function in the `coda` package in R:

$$\widehat{R} = \sqrt{\frac{d+3}{d+1}\frac{\widehat{V}}{W}}$$

where

$$d = 2\frac{\widehat{V}^2}{\mathrm{Var}(\widehat{V})}, \qquad \widehat{V} = \widehat{\sigma}^2 + \frac{B}{IJ}, \qquad \widehat{\sigma}^2 = \left(1 - \frac{1}{I}\right)W + \frac{B}{I}$$

5.2 Deviance Information Criterion

The deviance information criterion (DIC) is a model selection heuristic for hierarchical models much like the Akaike information criterion, AIC, and the Bayesian information criterion, BIC. As with AIC and BIC, given a set of models for $\boldsymbol{y}$, the one with the minimum DIC is preferred. DIC is based on the deviance,

$$D(\boldsymbol{y}, \boldsymbol{\eta}) = -2 \log p(\boldsymbol{y} \mid \boldsymbol{\eta})$$

where $\boldsymbol{y}$ is the data and $\boldsymbol{\eta}$ is the collection of model parameters. DIC itself is

$$\mathrm{DIC} = 2E(D(\boldsymbol{y}, \boldsymbol{\eta}) \mid \boldsymbol{y}) - D(\boldsymbol{y}, \widehat{\boldsymbol{\eta}})$$

where $\widehat{\boldsymbol{\eta}}$ is a suitable point estimate of $\boldsymbol{\eta}$. If $\boldsymbol{\eta}_i$ is the collection of parameter estimates of iteration i of the chain and $\overline{\boldsymbol{\eta}}$ is the collection of within-chain parameter means, then we can estimate DIC by

$$\begin{aligned}
\widehat{\mathrm{DIC}} &= \sum_{i=1}^{I} [2D(\boldsymbol{y} \mid \boldsymbol{\eta}_i)] - D(\boldsymbol{y}, \widehat{\boldsymbol{\eta}}) \\
&= -4 \sum_{i=1}^{I} \log p(\boldsymbol{y} \mid \boldsymbol{\eta}_i) + 2 \log p(\boldsymbol{y} \mid \overline{\boldsymbol{\eta}})
\end{aligned}$$

All that remains is to find $\log p(\boldsymbol{y} \mid \boldsymbol{\eta})$ for a given set of parameters, $\boldsymbol{\eta}$. Let $\lambda_{g,n} = \exp(\rho_n + \varepsilon_{g,n} + \eta(g, n))$, where

$$\eta(g, n) = \begin{cases} \phi_g - \alpha_g & \text{library } n \text{ is in treatment group 1} \\ \phi_g + \delta_g & \text{library } n \text{ is in treatment group 2} \\ \phi_g + \alpha_g & \text{library } n \text{ is in treatment group 3} \end{cases}$$

$$\begin{aligned}
\log p(\boldsymbol{y} \mid \boldsymbol{\eta}) &= \log \prod_{n=1}^{N} \prod_{g=1}^{G} \text{Poisson}(y_{g,n} \mid \lambda_{g,n}) \\
&= \sum_{n,g} \log \text{Poisson}(y_{g,n} \mid \lambda_{g,n}) \\
&= \sum_{n,g} \log \left(\frac{\exp(-\lambda_{g,n}) \lambda_{g,n}^{y_{g,n}}}{y_{g,n}!} \right) \\
&= \sum_{n,g} (-\lambda_{g,n} + y_{g,n} \log \lambda_{g,n} - \log(y_{g,n}!))
\end{aligned}$$

Given the size of the data, calculating $\sum_{n,g} - \log(y_{g,n}!)$ is intractable. Hence, in practice, we use

$$\text{DIC} = -4 \sum_{i=1}^{I} L(\boldsymbol{y} \mid \boldsymbol{\eta}_i) + 2L(\boldsymbol{y} \mid \overline{\boldsymbol{\eta}})$$

where

$$L(\boldsymbol{y}, \boldsymbol{\eta}) = \sum_{n,g} (-\lambda_{g,n} + y_{g,n} \log \lambda_{g,n}).$$

This approach is reasonable because removing the $- \log(y_{g,n}!)$ term inside the sum merely offsets the DIC values of all the models under comparison by the same constant.

A Derivations of the Full Conditionals

Recall:

- $k(n)$ = treatment group of library n.
- $\lambda_{g,n} = \exp(\rho_n + \varepsilon_{g,n} + \eta(g, n))$
- G_α = number of genes for which $\alpha_g \neq 0$
- G_δ = number of genes for which $\delta_g \neq 0$

Then from the model in Section 2, we get:

$$
\begin{aligned}
p(\nu \mid \cdots) &\propto \left[\prod_{g=1}^{G} \text{Inv-Gamma}\left(\gamma_g^2 \;\middle|\; \text{shape} = \frac{\nu}{2}, \text{scale} = \frac{\nu \cdot \tau^2}{2} \right) \right] \cdot \mathrm{U}(\nu \mid 0, d) \\
p(\rho_n \mid \cdots) &\propto \left[\prod_{g=1}^{G} \text{Poisson}(y_{g,n} \mid \exp(\rho_n + \varepsilon_{g,n} + \eta(g,n))) \right] \cdot \mathrm{N}(\rho_n \mid 0, \sigma_\rho^2) \\
p(\phi_g \mid \cdots) &\propto \left[\prod_{n=1}^{N} \text{Poisson}(y_{g,n} \mid \exp(\rho_n + \varepsilon_{g,n} + \eta(g,n))) \right] \cdot \mathrm{N}(\phi_g \mid \theta_\phi, \sigma_\phi^2) \\
p(\alpha_g \mid \cdots) &\propto \left[\prod_{k(n) \neq 2} \text{Poisson}(y_{g,n} \mid \exp(\rho_n + \varepsilon_{g,n} + \eta(g,n))) \right] \cdot \mathrm{N}(\alpha_g \mid \theta_\alpha, \sigma_\alpha^2) \\
p(\delta_g \mid \cdots) &\propto \left[\prod_{k(n) = 2} \text{Poisson}(y_{g,n} \mid \exp(\rho_n + \varepsilon_{g,n} + \eta(g,n))) \right] \cdot \mathrm{N}(\delta_g \mid \theta_\delta, \sigma_\delta^2) \\
p(\varepsilon_{g,n} \mid \cdots) &\propto \text{Poisson}(y_{g,n} \mid \exp(\rho_n + \varepsilon_{g,n} + \eta(g,n))) \cdot \mathrm{N}(\varepsilon_{g,n} \mid 0, \gamma_g^2)
\end{aligned}
$$

$$
\begin{aligned}
p\left(\sigma_\rho \mid \cdots\right) &= \left[\prod_{n=1}^{N} \mathrm{N}(\rho_n \mid 0, \sigma_\rho^2)\right] \cdot \mathrm{U}(\sigma_\rho \mid 0, s_\rho) \\
p(\gamma_g^2 \mid \cdots) &\propto \left[\prod_{n=1}^{N} \mathrm{N}(\varepsilon_{g,n} \mid 0, \gamma_g^2)\right] \cdot \text{Inv-Gamma}\left(\gamma_g^2 \;\middle|\; \text{shape} = \frac{\nu}{2}, \text{scale} = \frac{\nu \cdot \tau^2}{2}\right) \\
p(\tau^2 \mid \cdots) &\propto \left[\prod_{g=1}^{G} \text{Inv-Gamma}\left(\gamma_g^2 \;\middle|\; \text{shape} = \frac{\nu}{2}, \text{scale} = \frac{\nu \cdot \tau^2}{2}\right)\right] \cdot \text{Gamma}(\tau^2 \mid \text{shape} = a, \text{rate} = b) \\
p(\theta_\phi \mid \cdots) &\propto \left[\prod_{g=1}^{G} \mathrm{N}(\phi_g \mid \theta_\phi, \sigma_\phi^2)\right] \cdot \mathrm{N}(\theta_\phi \mid 0, c_\phi^2) \\
p(\theta_\alpha \mid \cdots) &\propto \left[\prod_{g=1}^{G} \mathrm{N}(\alpha_g \mid \theta_\alpha, \sigma_\alpha^2)\right] \cdot \mathrm{N}(\theta_\alpha \mid 0, c_\alpha^2) \\
p(\theta_\delta \mid \cdots) &\propto \left[\prod_{g=1}^{G} \mathrm{N}(\delta_g \mid \theta_\delta, \sigma_\delta^2)\right] \cdot \mathrm{N}(\theta_\delta \mid 0, c_\delta^2) \\
p(\sigma_\phi \mid \cdots) &\propto \left[\prod_{g=1}^{G} \mathrm{N}(\phi_g \mid \theta_\phi, \sigma_\phi^2)\right] \cdot \mathrm{U}(\sigma_\phi \mid 0, s_\phi) \\
p(\sigma_\alpha \mid \cdots) &\propto \left[\prod_{g=1}^{G} \mathrm{N}(\alpha_g \mid \theta_\alpha, \sigma_\alpha^2)\right] \cdot \mathrm{U}(\sigma_\alpha \mid 0, s_\alpha) \\
p(\sigma_\delta \mid \cdots) &\propto \left[\prod_{g=1}^{G} \mathrm{N}(\delta_g \mid \theta_\delta, \sigma_\delta^2)\right] \cdot \mathrm{U}(\sigma_\delta \mid 0, s_\delta)
\end{aligned}
$$

A.1 Transformations of Standard Deviations

Let σ be a standard deviation parameter and let $p(\sigma \mid \cdots)$ be its full conditional distribution. Then, by a transformation of variables,

$$
\begin{aligned}
p(\sigma^2 \mid \cdots) &= p(\sqrt{\sigma^2} \mid \cdots) \cdot \left| \frac{d}{d\sigma^2} \sqrt{\sigma^2} \right| \\
&= p(\sigma \mid \cdots) \frac{1}{2} (\sigma^2)^{-1/2}
\end{aligned}
$$

I use this transformation several times in the next sections.

A.2 $p(\nu \mid \cdots)$: Metropolis

$$
\begin{aligned}
p(\nu \mid \cdots) &= \left[\prod_{g=1}^{G} \text{Inv-Gamma}\left(\gamma_g^2 \mid \text{shape} = \frac{\nu}{2},\ \text{scale} = \frac{\nu \cdot \tau^2}{2}\right)\right] \cdot \text{U}(\nu \mid 0, d) \\
&\propto \prod_{g=1}^{G}\left[\Gamma\left(\nu/2\right)^{-1}\left(\frac{\nu \cdot \tau^2}{2}\right)^{\nu/2}\left(\gamma_g^2\right)^{-(\nu/2+1)} \exp\left(-\frac{1}{\gamma_g^2}\frac{\nu \cdot \tau^2}{2}\right)\right] I(0 < \nu < d) \\
&\propto \Gamma\left(\nu/2\right)^{-G}\left(\frac{\nu \cdot \tau^2}{2}\right)^{G\nu/2}\left(\prod_{g=1}^{G}\gamma_g^2\right)^{-(\nu/2+1)} \exp\left(-\frac{\nu \cdot \tau^2}{2}\sum_{g=1}^{G}\frac{1}{\gamma_g^2}\right) I(0 < \nu < d)
\end{aligned}
$$

A.3 $p(\rho_n \mid \cdots)$: Metropolis

$$
\begin{aligned}
p(\rho_n \mid \cdots) &\propto \left[\prod_{g=1}^{G} \text{Poisson}(y_{g,n} \mid \lambda_{g,n})\right] \cdot \text{N}(\rho_n \mid 0, \sigma_\rho^2) \\
&\propto \left[\prod_{g=1}^{G} \lambda_{g,n}^{y_{g,n}} \exp(-\lambda_{g,n})\right] \exp\left(-\frac{\rho_n^2}{2\sigma_\rho^2}\right) \\
&= \exp\left(\sum_{g=1}^{G}\left[y_{g,n}\log\lambda_{g,n} - \lambda_{g,n}\right] - \frac{\rho_n^2}{2\sigma_\rho^2}\right) \\
&= \exp\left(\sum_{g=1}^{G}\left[y_{g,n}(\rho_n + \varepsilon_{g,n} + \eta(g,n)) - \exp(\rho_n + \varepsilon_{g,n} + \eta(g,n))\right] - \frac{\rho_n^2}{2\sigma_\rho^2}\right) \\
&= \exp\left(\rho_n G\overline{y}_{.n} + \sum_{g=1}^{G}\left[y_{g,n}(\varepsilon_{g,n} + \eta(g,n))\right] - \sum_{g=1}^{G}\exp(\rho_n + \varepsilon_{g,n} + \eta(g,n)) - \frac{\rho_n^2}{2\sigma_\rho^2}\right) \\
&\propto \exp\left(\rho_n G\overline{y}_{.n} - \exp(\rho_n)\sum_{g=1}^{G}\exp(\varepsilon_{g,n} + \eta(g,n)) - \frac{\rho_n^2}{2\sigma_\rho^2}\right) \\
&\propto \exp\left(\rho_n G\overline{y}_{.n} - \frac{\rho_n^2}{2\sigma_\rho^2} - \exp(\rho_n)\sum_{g=1}^{G}\exp(\varepsilon_{g,n} + \eta(g,n))\right)
\end{aligned}
$$

A.4 $p(\phi_g \mid \cdots)$: Metropolis

$$\begin{aligned}
p(\phi_g \mid \cdots) &= \left[\prod_{n=1}^{N} \text{Poisson}(y_{g,n} \mid \lambda_{g,n})\right] \cdot \text{N}(\phi_g \mid \theta_\phi, \sigma_\phi^2) \\
&\propto \left[\prod_{n=1}^{N} \lambda_{g,n}^{y_{g,n}} \exp(-\lambda_{g,n})\right] \cdot \exp\left(-\frac{(\phi_g - \theta_\phi)^2}{2\sigma_\phi^2}\right) \\
&= \exp\left(\sum_{n=1}^{N} [y_{g,n} \log \lambda_{g,n} - \lambda_{g,n}] - \frac{(\phi_g - \theta_\phi)^2}{2\sigma_\phi^2}\right) \\
&= \exp\left(\sum_{n=1}^{N} [y_{g,n}(\rho_n + \varepsilon_{g,n} + \eta(g,n)) - \exp(\rho_n + \varepsilon_{g,n} + \eta(g,n))] - \frac{(\phi_g - \theta_\phi)^2}{2\sigma_\phi^2}\right) \\
&\propto \exp\left(\sum_{n=1}^{N} [y_{g,n}\eta(g,n) - \exp(\rho_n + \varepsilon_{g,n} + \eta(g,n))] - \frac{(\phi_g - \theta_\phi)^2}{2\sigma_\phi^2}\right) \\
&\propto \exp\left(\sum_{n=1}^{N} [y_{g,n}\eta(g,n)] - \sum_{n=1}^{N} [\exp(\rho_n + \varepsilon_{g,n} + \eta(g,n))] - \frac{(\phi_g - \theta_\phi)^2}{2\sigma_\phi^2}\right) \\
&= \exp\left(\sum_{n=1}^{N} [y_{g,n}\eta(g,n)] - \frac{(\phi_g - \theta_\phi)^2}{2\sigma_\phi^2} - \sum_{n=1}^{N} [\exp(\rho_n + \varepsilon_{g,n} + \eta(g,n))]\right)
\end{aligned}$$

and

$$\begin{aligned}
\sum_{n=1}^{N} [y_{g,n}\eta(g,n)] &= \sum_{k(n)=1} [y_{g,n}\eta(g,n)] + \sum_{k(n)=2} [y_{g,n}\eta(g,n)] + \sum_{k(n)=3} [y_{g,n}\eta(g,n)] \\
&= \sum_{k(n)=1} [y_{g,n}(\phi_g - \alpha_g)] + \sum_{k(n)=2} [y_{g,n}(\phi_g + \delta_g)] + \sum_{k(n)=3} [y_{g,n}(\phi_g + \alpha_g)] \\
&= \phi_g N \overline{y}_{g.} + \text{ constant}
\end{aligned}$$

and

$$
\begin{aligned}
\sum_{n=1}^{N}[\exp(\rho_n+\varepsilon_{g,n}+\eta(g,n))] &= \sum_{k(n)=1}[\exp(\rho_n+\varepsilon_{g,n}+\eta(g,n))]+\sum_{k(n)=2}[\exp(\rho_n+\varepsilon_{g,n}+\eta(g,n))] \\
&\quad+\sum_{k(n)=3}[\exp(\rho_n+\varepsilon_{g,n}+\eta(g,n))] \\
&= \sum_{k(n)=1}[\exp(\rho_n+\varepsilon_{g,n}+\phi_g-\alpha_g)]+\sum_{k(n)=2}[\exp(\rho_n+\varepsilon_{g,n}+\phi_g+\delta_g)] \\
&\quad+\sum_{k(n)=3}[\exp(\rho_n+\varepsilon_{g,n}+\phi_g+\alpha_g)] \\
&= \exp(\phi_g)\left[\sum_{k(n)=1}[\exp(\rho_n+\varepsilon_{g,n}-\alpha_g)]+\sum_{k(n)=2}[\exp(\rho_n+\varepsilon_{g,n}+\delta_g)]\right. \\
&\quad\left.+\sum_{k(n)=3}[\exp(\rho_n+\varepsilon_{g,n}+\alpha_g)]\right] \\
&= \exp(\phi_g)\left[\exp(-\alpha_g)\sum_{k(n)=1}[\exp(\rho_n+\varepsilon_{g,n})]+\exp(\delta_g)\sum_{k(n)=2}[\exp(\rho_n+\varepsilon_{g,n})]\right. \\
&\quad\left.+\exp(\alpha_g)\sum_{k(n)=3}[\exp(\rho_n+\varepsilon_{g,n})]\right]
\end{aligned}
$$

so

$$
\begin{aligned}
p(\phi_g \mid \cdots) \propto \exp\Bigg(&\phi_g N\overline{y}_{g.}-\frac{(\phi_g-\theta_\phi)^2}{2\sigma_\phi^2}-\exp(\phi_g)\Bigg[\exp(-\alpha_g)\sum_{k(n)=1}[\exp(\rho_n+\varepsilon_{g,n})] \\
&\exp(\delta_g)\sum_{k(n)=2}[\exp(\rho_n+\varepsilon_{g,n})]+\exp(\alpha_g)\sum_{k(n)=3}[\exp(\rho_n+\varepsilon_{g,n})]\Bigg]\Bigg)
\end{aligned}
$$

A.5 $p(\alpha_g \mid \cdots)$: Metropolis

Similar to ϕ_g,

$$
p(\alpha_g \mid \cdots) \propto \exp\left(\sum_{k(n)\neq 2}[y_{g,n}\eta(g,n)]-\frac{(\alpha_g-\theta_\alpha)^2}{2\sigma_\alpha^2}-\sum_{k(n)\neq 2}[\exp(\rho_n+\varepsilon_{g,n}+\eta(g,n))]\right)
$$

and

$$\begin{aligned}
\sum_{k(n)\neq 2} [y_{g,n}\eta(g,n)] &= \sum_{k(n)=1} [y_{g,n}\eta(g,n)] + \sum_{k(n)=3} [y_{g,n}\eta(g,n)] \\
&= \sum_{k(n)=1} [y_{g,n}(\phi_g - \alpha_g)] + \sum_{k(n)=3} [y_{g,n}(\phi_g + \alpha_g)] \\
&= \alpha_g \left(\sum_{k(n)=3} y_{g,n} - \sum_{k(n)=1} y_{g,n} \right) + \text{ constant}
\end{aligned}$$

and

$$\begin{aligned}
\sum_{k(n)\neq 2} \exp(\rho_n + \varepsilon_{g,n} + \eta(g,n)) &= \sum_{k(n)=1} [\exp(\rho_n + \varepsilon_{g,n} + \eta(g,n))] + \sum_{k(n)=3} [\exp(\rho_n + \varepsilon_{g,n} + \eta(g,n))] \\
&= \sum_{k(n)=1} [\exp(\rho_n + \varepsilon_{g,n} + \phi_g - \alpha_g)] + \sum_{k(n)=3} [\exp(\rho_n + \varepsilon_{g,n} + \phi_g + \alpha_g)] \\
&= \exp(-\alpha_g)\exp(\phi_g) \sum_{k(n)=1} [\exp(\rho_n + \varepsilon_{g,n})] + \exp(\alpha_g)\exp(\phi_g) \sum_{k(n)=3} [\exp(\rho_n + \varepsilon_{g,n})]
\end{aligned}$$

so

$$\begin{aligned}
p(\alpha_g \mid \cdots) \propto \exp \Bigg(&\alpha_g \left(\sum_{k(n)=3} y_{g,n} - \sum_{k(n)=1} y_{g,n} \right) - \frac{(\alpha_g - \theta_\alpha)^2}{2\sigma_\alpha^2} - \exp(-\alpha_g)\exp(\phi_g) \sum_{k(n)=1} [\exp(\rho_n + \varepsilon_{g,n})] \\
&- \exp(\alpha_g)\exp(\phi_g) \sum_{k(n)=3} [\exp(\rho_n + \varepsilon_{g,n})] \Bigg)
\end{aligned}$$

A.6 $p(\delta_g \mid \cdots)$: Metropolis

Similar to ϕ_g,

$$p(\delta_g \mid \cdots) \propto \exp \left(\sum_{k(n)=2} [y_{g,n}\eta(g,n)] - \frac{(\delta_g - \theta_\delta)^2}{2\sigma_\delta^2} - \sum_{k(n)\neq 2} [\exp(\rho_n + \varepsilon_{g,n} + \eta(g,n))] \right)$$

and

$$\begin{aligned}\sum_{k(n)=2}[y_{g,n}\eta(g,n)] &= \sum_{k(n)=2}[y_{g,n}(\phi_g+\delta_g)] \\ &= \delta_g \sum_{k(n)=2} y_{g,n} + \text{ constant}\end{aligned}$$

and

$$\begin{aligned}\sum_{k(n)\neq 2}\exp(\rho_n+\varepsilon_{g,n}+\eta(g,n)) &= \sum_{k(n)\neq 2}\exp(\rho_n+\varepsilon_{g,n}+\phi_g+\delta_g) \\ &= \exp(\delta_g)\exp(\phi_g)\sum_{k(n)\neq 2}\exp(\rho_n+\varepsilon_{g,n})\end{aligned}$$

so

$$p(\delta_g \mid \cdots) \propto \exp\left(\delta_g \sum_{k(n)=2} y_{g,n} - \frac{(\delta_g-\theta_\delta)^2}{2\sigma_\delta^2} - \exp(\delta_g)\exp(\phi_g)\sum_{k(n)\neq 2}\exp(\rho_n+\varepsilon_{g,n})\right)$$

A.7 $p(\varepsilon_{g,n} \mid \cdots)$ Metropolis

$$\begin{aligned}p(\varepsilon_{g,n} \mid \cdots) &= \text{Poisson}(y_{g,n} \mid \lambda_{g,n}) \cdot \text{N}(\varepsilon_{g,n} \mid 0, \gamma_g^2) \\ &\propto \lambda_{g,n}^{y_{g,n}} \exp(-\lambda_{g,n}) \exp\left(-\frac{\varepsilon_{g,n}^2}{2\gamma_g^2}\right) \\ &= \exp\left(y_{g,n}\log\lambda_{g,n} - \lambda_{g,n} - \frac{\varepsilon_{g,n}^2}{2\gamma_g^2}\right) \\ &= \exp\left(y_{g,n}(\rho_n+\varepsilon_{g,n}+\eta(g,n)) - \exp(\rho_n+\varepsilon_{g,n}+\eta(g,n)) - \frac{\varepsilon_{g,n}^2}{2\gamma_g^2}\right) \\ &= \exp\left(y_{g,n}\varepsilon_{g,n} - \exp(\rho_n+\varepsilon_{g,n}+\eta(g,n)) - \frac{\varepsilon_{g,n}^2}{2\gamma_g^2}\right) \\ &= \exp\left(y_{g,n}\varepsilon_{g,n} - \frac{\varepsilon_{g,n}^2}{2\gamma_g^2} - \exp(\varepsilon_{g,n})\exp(\rho_n+\eta(g,n))\right)\end{aligned}$$

A.8 $p(\theta_\phi \mid \cdots)$: Normal

$$
\begin{aligned}
p(\theta_\phi \mid \cdots) &= \left[\prod_{g=1}^{G} \mathrm{N}(\phi_g \mid \theta_\phi, \sigma_\phi^2)\right] \cdot \mathrm{N}(\theta_\phi \mid 0, c_\phi^2) \\
&\propto \left[\prod_{g=1}^{G} \exp\left(-\frac{(\phi_g - \theta_\phi)^2}{2\sigma_\phi^2}\right)\right] \exp\left(-\frac{\theta_\phi^2}{2c_\phi^2}\right) \\
&= \exp\left(-\sum_{g=1}^{G} \frac{(\phi_g - \theta_\phi)^2}{2\sigma_\phi^2}\right) \exp\left(-\frac{\theta_\phi^2}{2c_\phi^2}\right) \\
&= \exp\left(-\frac{\sum_{g=1}^{G} \phi_g^2 - 2\theta_\phi \sum_{g=1}^{G} \phi_g + G\theta_\phi^2}{2\sigma_\phi^2}\right) \exp\left(-\frac{\theta_\phi^2}{2c_\phi^2}\right) \\
&= \exp\left(-\frac{\sum_{g=1}^{G} \phi_g^2 - 2\theta_\phi \sum_{g=1}^{G} \phi_g + G\theta_\phi^2}{2\sigma_\phi^2} - \frac{\theta_\phi^2}{2c_\phi^2}\right) \\
&= \exp\left(-\frac{c_\phi^2 \sum_{g=1}^{G} \phi_g^2 - 2c_\phi^2 (\sum_{g=1}^{G} \phi_g)\theta_\phi + Gc_\phi^2\theta_\phi^2}{2\sigma_\phi^2 c_\phi^2} - \frac{\sigma_\phi^2\theta_\phi^2}{2\sigma_\phi^2 c_\phi^2}\right) \\
&= \exp\left(-\frac{c_\phi^2 \sum_{g=1}^{G} \phi_g^2 - 2c_\phi^2 (\sum_{g=1}^{G} \phi_g)\theta_\phi + (Gc_\phi^2 + \sigma_\phi^2)\theta_\phi^2}{2\sigma_\phi^2 c_\phi^2}\right) \\
&\propto \exp\left(-\frac{(Gc_\phi^2 + \sigma_\phi^2)\left(\theta_\phi - \frac{c_\phi^2(\sum_{g=1}^{G}\phi_g)}{Gc_\phi^2 + \sigma_\phi^2}\right)^2}{2\sigma_\phi^2 c_\phi^2}\right)
\end{aligned}
$$

Hence:

$$
p(\theta_\phi \mid \cdots) = \mathrm{N}\left(\theta_\phi \;\middle|\; \frac{c_\phi^2 \sum_{g=1}^{G} \phi_g}{Gc_\phi^2 + \sigma_\phi^2},\ \frac{c_\phi^2 \sigma_\phi^2}{Gc_\phi^2 + \sigma_\phi^2}\right)
$$

A.9 $p\left(\theta_\alpha \mid \cdots\right)$: Normal

$$
p(\theta_\alpha \mid \cdots) \propto \left[\prod_{g=1}^{G} \mathrm{N}(\alpha_g \mid \theta_\alpha, \sigma_\alpha^2)]\right] \cdot \mathrm{N}(\theta_\alpha \mid 0, c_\alpha^2)
$$

From algebra similar to the derivation of $p(\theta_\phi \mid \cdots)$,

$$
p(\theta_\alpha \mid \cdots) = N\left(\theta_\alpha \;\middle|\; \frac{c_\alpha^2 \sum_{g=1}^{G} \alpha_g}{G_\alpha c_\alpha^2 + \sigma_\alpha^2},\ \frac{c_\alpha^2 \sigma_\alpha^2}{G_\alpha c_\alpha^2 + \sigma_\alpha^2}\right)
$$

A.10 $p(\theta_\delta \mid \cdots)$: Normal

$$p(\theta_\delta \mid \cdots) \propto \left[\prod_{g=1}^{G} \mathrm{N}(\delta_g \mid \theta_\delta, \sigma_\delta^2)]\right] \cdot \mathrm{N}(\theta_\delta \mid 0, c_\delta^2)$$

From algebra similar to the derivation of $p(\theta_\phi \mid \cdots)$,

$$p(\theta_\delta \mid \cdots) = N\left(\theta_\delta \;\middle|\; \frac{c_\delta^2 \sum_{g=1}^{G} \delta_g}{G_\delta c_\delta^2 + \sigma_\delta^2}, \;\; \frac{c_\delta^2 \sigma_\delta^2}{G_\delta c_\delta^2 + \sigma_\delta^2}\right)$$

A.11 $p\left(\frac{1}{\sigma_\phi^2} \mid \ldots\right)$: Truncated Gamma

$$\begin{aligned}
p(\sigma_\phi^2 \mid \cdots) &= p(\sigma_\phi \mid \cdots)\frac{1}{2}(\sigma_\phi^2)^{-1/2} \qquad \text{(transformation in Section A.1)} \\
&\propto \left[\prod_{g=1}^{G} \mathrm{N}(\phi_g \mid \theta_\phi, \sigma_\phi^2)\right] \cdot \mathrm{U}(\sigma_\phi \mid 0, s_\phi)(\sigma_\phi^2)^{-1/2} \\
&\propto \left[\prod_{g=1}^{G} (\sigma_\phi^2)^{-1/2} \exp\left(-\frac{(\phi_g - \theta_\phi)^2}{2\sigma_\phi^2}\right)\right] \mathrm{I}(0 < \sigma_\phi^2 < s_\phi^2)(\sigma_\phi^2)^{-1/2} \\
&= (\sigma_\phi^2)^{-G/2} \exp\left(-\sum_{g=1}^{G} \frac{(\phi_g - \theta_\phi)^2}{2\sigma_\phi^2}\right) \mathrm{I}(0 < \sigma_\phi^2 < s_\phi^2)(\sigma_\phi^2)^{-1/2} \\
&= (\sigma_\phi^2)^{-(G/2-1/2+1)} \exp\left(-\frac{1}{\sigma_\phi^2}\frac{1}{2}\sum_{g=1}^{G}(\phi_g - \theta_\phi)^2\right) \mathrm{I}(0 < \sigma_\phi^2 < s_\phi^2)
\end{aligned}$$

which is the kernel of a truncated inverse gamma distribution. Hence:

$$p\left(\frac{1}{\sigma_\phi^2} \;\middle|\; \cdots\right) = \mathrm{Gamma}\left(\text{shape} = \frac{G-1}{2}, \ \text{rate} = \frac{1}{2}\sum_{g=1}^{G}(\phi_g - \theta_\phi)^2\right) \mathrm{I}\left(\frac{1}{\sigma_\phi^2} > \frac{1}{s_\phi^2}\right)$$

A.12 $p\left(\frac{1}{\sigma_\alpha^2} \mid \cdots\right)$: Truncated Gamma

Analogously to σ_ϕ,

$$p\left(\frac{1}{\sigma_\alpha^2} \;\middle|\; \cdots\right) = \mathrm{Gamma}\left(\text{shape} = \frac{G-1}{2}, \ \text{rate} = \frac{1}{2}\sum_{g=1}^{G}(\alpha_g - \theta_\alpha)^2\right) \mathrm{I}\left(\frac{1}{\sigma_\alpha^2} > \frac{1}{s_\alpha^2}\right)$$

A.13 $p\left(\frac{1}{\sigma_\delta^2} \mid \cdots\right)$: Truncated Gamma

Analogously to σ_ϕ,

$$p\left(\frac{1}{\sigma_\delta^2} \;\middle|\; \cdots\right) = \text{Gamma}\left(\text{shape} = \frac{G-1}{2},\ \text{rate} = \frac{1}{2}\sum_{g=1}^{G}(\delta_g - \theta_\delta)^2\right) \mathrm{I}\left(\frac{1}{\sigma_\delta^2} > \frac{1}{s_\delta^2}\right)$$

A.14 $p\left(\frac{1}{\sigma_\rho^2} \mid \cdots\right)$ Truncated Gamma

$$\begin{aligned}
p(\sigma_\rho^2 \mid \cdots) &= p(\sigma_\rho \mid \cdots)\frac{1}{2}(\sigma_\rho^2)^{-1/2} \qquad \text{(transformation in Section A.1)} \\
&\propto \left[\prod_{n=1}^{N} \mathrm{N}(\rho_n \mid 0, \sigma_\rho^2)\right] \cdot \mathrm{U}(\sigma_\rho \mid 0, s_\rho)\frac{1}{2}(\sigma_\rho^2)^{-1/2} \\
&\propto \prod_{n=1}^{N}\left[\frac{1}{\sqrt{\sigma_\rho^2}}\exp\left(-\frac{\rho_n^2}{2\sigma_\rho^2}\right)\right] \cdot \mathrm{I}(0 < \sigma_\rho < s_\rho)(\sigma_\rho^2)^{-1/2} \\
&= (\sigma_\rho^2)^{-N/2}\exp\left(-\frac{1}{\sigma_\rho^2}\frac{1}{2}\sum_{n=1}^{N}\rho_n^2\right) \cdot \mathrm{I}(0 < \sigma_\rho < s_\rho)(\sigma_\rho^2)^{-1/2} \\
&= (\sigma_\rho^2)^{-(N/2-1/2+1)}\exp\left(-\frac{1}{\sigma_\rho^2}\frac{1}{2}\sum_{n=1}^{N}\rho_n^2\right) \cdot \mathrm{I}(0 < \sigma_\rho < s_\rho)
\end{aligned}$$

which is the kernel of a truncated inverse gamma distribution. Hence:

$$p\left(\frac{1}{\sigma_\rho^2} \mid \cdots\right) = \text{Gamma}\left(\frac{1}{\sigma_\rho^2} \mid \text{shape} = \frac{N-1}{2},\ \text{rate} = \frac{1}{2}\sum_{n=1}^{N}\rho_n^2\right) I\left(\frac{1}{\sigma_\rho^2} > \frac{1}{s_\rho^2}\right)$$

A.15 $p(\tau^2 \mid \cdots)$: Gamma

$$
\begin{aligned}
p(\tau^2 \mid \cdots) &= \left[\prod_{g=1}^{G} \text{Inv-Gamma}\left(\gamma_g^2 \mid \text{shape} = \frac{\nu}{2}, \text{scale} = \frac{\nu \cdot \tau^2}{2}\right)\right] \cdot \text{Gamma}(\tau^2 \mid \text{shape} = a, \text{rate} = b) \\
&\propto \left[\Gamma\left(\nu/2\right)^{-G} \left(\frac{\nu \cdot \tau^2}{2}\right)^{G\nu/2} \left(\prod_{g=1}^{G} \gamma_g^2\right)^{-(\nu/2+1)} \exp\left(-\frac{\nu \cdot \tau^2}{2} \sum_{g=1}^{G} \frac{1}{\gamma_g^2}\right)\right] \cdot (\tau^2)^{a-1} \exp\left(-b\tau^2\right) \\
&\propto \left[\left(\tau^2\right)^{G\nu/2} \exp\left(-\tau^2 \cdot \frac{\nu}{2} \sum_{g=1}^{G} \frac{1}{\gamma_g^2}\right)\right] \cdot (\tau^2)^{a-1} \exp\left(-b\tau^2\right) \\
&= (\tau^2)^{G\nu/2+a-1} \exp\left(-\tau^2 \left(b + \frac{\nu}{2} \sum_{g=1}^{G} \frac{1}{\gamma_g^2}\right)\right)
\end{aligned}
$$

Hence:

$$
p(\tau^2 \mid \cdots) = \text{Gamma}\left(\tau^2 \mid \text{shape} = a + \frac{G\nu}{2},\ \text{rate} = b + \frac{\nu}{2} \sum_{g=1}^{G} \frac{1}{\gamma_g^2}\right)
$$

A.16 $p\left(\frac{1}{\gamma_g^2} \mid \cdots\right)$ Gamma

$$
\begin{aligned}
p(\gamma_g^2 \mid \cdots) &= \left[\prod_{n=1}^{N} \text{N}(\varepsilon_{g,n} \mid 0, \gamma_g^2)\right] \cdot \text{Inv-Gamma}\left(\gamma_g^2 \mid \text{shape} = \frac{\nu}{2}, \text{scale} = \frac{\nu \cdot \tau^2}{2}\right) \\
&\propto \left[\prod_{n=1}^{N} (\gamma_g^2)^{-1/2} \exp\left(-\frac{1}{\gamma_g^2} \frac{\varepsilon_{g,n}^2}{2}\right)\right] \cdot \left(\gamma_g^2\right)^{-(\nu/2+1)} \exp\left(-\frac{1}{\gamma_g^2} \frac{\nu \cdot \tau^2}{2}\right) \\
&= \left[(\gamma_g^2)^{-N/2} \exp\left(-\frac{1}{\gamma_g^2} \frac{1}{2} \sum_{n=1}^{N} \varepsilon_{g,n}^2\right)\right] \cdot \left(\gamma_g^2\right)^{-(\nu/2+1)} \exp\left(-\frac{1}{\gamma_g^2} \frac{\nu \cdot \tau^2}{2}\right) \\
&= (\gamma_g^2)^{-((N+\nu)/2+1)} \exp\left(-\frac{1}{\gamma_g^2} \frac{1}{2} \left(\nu \cdot \tau^2 + \sum_{n=1}^{N} \varepsilon_{g,n}^2\right)\right)
\end{aligned}
$$

which is the kernel of an inverse gamma distribution. Hence:

$$
p\left(\frac{1}{\gamma_g^2} \,\middle|\, \cdots\right) = \text{Gamma}\left(\frac{1}{\gamma_g^2} \,\middle|\, \text{shape} = \frac{N+\nu}{2},\ \text{rate} = \frac{1}{2}\left(\nu \cdot \tau^2 + \sum_{n=1}^{N} \varepsilon_{g,n}^2\right)\right)
$$

B Metropolis steps

B.1 A common proposal for ρ_n, ϕ_g, α_g, δ_g, and $\varepsilon_{g,n}$

The full conditionals of ρ_n, ϕ_g, α_g, δ_g, and $\varepsilon_{g,n}$ all have the form,

$$
\begin{aligned}
\log p(\theta \mid \cdots) &= A\theta + B(\theta - C)^2 + De^{\theta} + Ee^{-\theta} \\
&\approx A\theta + B(\theta - C)^2 + De^{\widehat{\theta}}\left(1 + \theta - \widehat{\theta} + \frac{(\theta - \widehat{\theta})^2}{2}\right) + Ee^{-\widehat{\theta}}\left(1 - \theta + \widehat{\theta} + \frac{(\theta - \widehat{\theta})^2}{2}\right) \\
&= BC^2 + De^{\widehat{\theta}} - De^{\widehat{\theta}}\widehat{\theta} + \frac{1}{2}De^{\widehat{\theta}}\widehat{\theta}^2 + Ee^{-\widehat{\theta}} + Ee^{-\widehat{\theta}}\widehat{\theta} + \frac{1}{2}Ee^{-\widehat{\theta}}\widehat{\theta}^2 \\
&\quad + A\theta - 2BC\theta + De^{\widehat{\theta}}\theta - De^{\widehat{\theta}}\widehat{\theta}\theta - Ee^{-\widehat{\theta}}\theta - Ee^{-\widehat{\theta}}\widehat{\theta}\theta \\
&\quad + B\theta^2 + \frac{1}{2}De^{\widehat{\theta}}\theta^2 + \frac{1}{2}Ee^{-\widehat{\theta}}\theta^2 \\
&= \left[BC^2 + De^{\widehat{\theta}} - De^{\widehat{\theta}}\widehat{\theta} + \frac{1}{2}De^{\widehat{\theta}}\widehat{\theta}^2 + Ee^{-\widehat{\theta}} + Ee^{-\widehat{\theta}}\widehat{\theta} + \frac{1}{2}Ee^{-\widehat{\theta}}\widehat{\theta}^2\right] \\
&\quad + \left[A - 2BC + De^{\widehat{\theta}} - De^{\widehat{\theta}}\widehat{\theta} - Ee^{-\widehat{\theta}} - Ee^{-\widehat{\theta}}\widehat{\theta}\right]\theta \\
&\quad + \left[B + \frac{1}{2}De^{\widehat{\theta}} + \frac{1}{2}Ee^{-\widehat{\theta}}\right]\theta^2 \\
&= \left[BC^2 + De^{\widehat{\theta}}\left(1 - \widehat{\theta} + \frac{1}{2}\widehat{\theta}^2\right) + Ee^{-\widehat{\theta}}\left(1 + \widehat{\theta} + \frac{1}{2}\widehat{\theta}^2\right)\right] \\
&\quad + \left[A - 2BC + De^{\widehat{\theta}}\left(1 - \widehat{\theta}\right) - Ee^{-\widehat{\theta}}\left(1 + Ee^{-\widehat{\theta}}\widehat{\theta}\right)\right]\theta \\
&\quad + \left[B + \frac{1}{2}De^{\widehat{\theta}} + \frac{1}{2}Ee^{-\widehat{\theta}}\right]\theta^2
\end{aligned}
$$

B.2 Calculating $\widehat{c}_n$

Let $g(\rho_n)$ be the kernel of the log full conditional density of ρ_n. Then,

$$
g(\rho_n) = \rho_n G\overline{y}_{.n} - \exp(\rho_n)\sum_{g=1}^{G}\exp(\varepsilon_{g,n} + \eta(g,n)) - \frac{\rho_n^2}{2\sigma_\rho^2}
$$

Differentiating,

$$
g'(\rho_n) = G\overline{y}_{.n} - \exp(\rho_n)\sum_{g=1}^{G}\exp(\varepsilon_{g,n} + \eta(g,n)) - \frac{\rho_n}{\sigma_\rho^2}
$$

We let $\widehat{c}_n$ be the root of this derivative.

$$
0 = G\overline{y}_{.n} - \exp(\widehat{c}_n)\sum_{g=1}^{G}\exp(\varepsilon_{g,n} + \eta(g,n)) - \frac{\widehat{c}_n}{\sigma_\rho^2}
$$

Using a quadratic approximation to the exponential function,

$$\begin{aligned}
0 &= G\overline{y}_{.n} - \left(1 + \widehat{c}_n + \frac{\widehat{c}_n^2}{2}\right) \underbrace{\sum_{g=1}^{G} \exp(\varepsilon_{g,n} + \eta(g,n))}_{S} - \frac{\widehat{c}_n}{\sigma_\rho^2} \\
&= (G\overline{y}_{.n} - S) + \left(-S - \frac{1}{\sigma_\rho^2}\right)\widehat{c}_n + \left(\frac{S}{2}\right)\widehat{c}_n^2
\end{aligned}$$

Using the quadratic formula, we get

$$\widehat{c}_n = \frac{S + \frac{1}{\sigma_\rho^2} \pm \sqrt{\left(S + \frac{1}{\sigma_\rho^2}\right)^2 - 2S(G\overline{y}_{.n} - S)}}{S}$$

In practice, I will use the root with the higher value of $g(\widehat{c}_n)$.

B.3 Calculating $\widehat{\varepsilon}_{g,n}$

Let $g(\varepsilon_{g,n})$ be the kernel of the log full conditional density of $\varepsilon_{g,n}$.

$$\begin{aligned}
g(\varepsilon_{g,n}) &= y_{g,n}\varepsilon_{g,n} - \exp(\rho_n + \varepsilon_{g,n} + \eta(g,n)) - \frac{\varepsilon_{g,n}^2}{2\gamma_g^2} \\
&= y_{g,n}\varepsilon_{g,n} - \exp(\varepsilon_{g,n})\exp(\rho_n + \eta(g,n)) - \frac{\varepsilon_{g,n}^2}{2\gamma_g^2}
\end{aligned}$$

Differentiating with respect to $\varepsilon_{g,n}$,

$$g(\varepsilon_{g,n}) = y_{g,n} - \exp(\varepsilon_{g,n})\exp(\rho_n + \eta(g,n)) - \frac{\varepsilon_{g,n}}{\gamma_g^2}$$

We let $\widehat{\varepsilon}_{g,n}$ be the root of this derivative.

$$0 = y_{g,n} - \exp(\widehat{\varepsilon}_{g,n})\exp(\rho_n + \eta(g,n)) - \frac{\widehat{\varepsilon}_{g,n}}{\gamma_g^2}$$

Taking the quadratic approximation to the exponential,

$$\begin{aligned}
0 &= y_{g,n} - \left(1 + \widehat{\varepsilon}_{g,n} + \frac{\widehat{\varepsilon}_{g,n}}{2}\right) \underbrace{\exp(\rho_n + \eta(g,n))}_{S} - \frac{\widehat{\varepsilon}_{g,n}}{\gamma_g^2} \\
&= (y_{g,n} - S) + \left(-S - \frac{1}{\gamma_g^2}\right)\widehat{\varepsilon}_{g,n} + \left(\frac{S}{2}\right)\widehat{\varepsilon}_{g,n}^2
\end{aligned}$$

Using the quadratic formula,

$$\widehat{\varepsilon}_{g,n} = \frac{\left(S + \frac{1}{\gamma_g^2}\right) \pm \sqrt{\left(S + \frac{1}{\gamma_g^2}\right)^2 - 2S(y_{g,n} - S)}}{S}$$

In practice, I will use the root with the higher value of $g(\widehat{\varepsilon}_{g,n})$.

B.4 Calculating $\widehat{\phi}_g$

Let $g(\phi_g)$ be the kernel of the log full conditional density of ϕ_g. Then,

$$
\begin{aligned}
g(\phi_g) &= \sum_{n=1}^{N} \left[y_{g,n}\eta(g,n) - \exp(\rho_n + \varepsilon_{g,n} - \eta(g,n)) \right] - \frac{(\phi_g - \theta_\phi)^2}{2{\sigma_\phi}^2} \\
&= \sum_{\text{group}(n)=1} \left[y_{g,n}(\phi_g - \alpha_g) - \exp(\rho_n + \varepsilon_{g,n} - (\phi_g - \alpha_g)) \right] \\
&\quad + \sum_{\text{group}(n)=2} \left[y_{g,n}(\phi_g + \delta_g) - \exp(\rho_n + \varepsilon_{g,n} - (\phi_g + \delta_g)) \right] \\
&\quad + \sum_{\text{group}(n)=3} \left[y_{g,n}(\phi_g + \alpha_g) - \exp(\rho_n + \varepsilon_{g,n} - (\phi_g + \alpha_g)) \right] \\
&\quad - \frac{(\phi_g - \theta_\phi)^2}{2\sigma_\phi^2} \\
&= \sum_{\text{group}(n)=1} y_{g,n}(\phi_g - \alpha_g) - \sum_{\text{group}(n)=1} \exp(\rho_n + \varepsilon_{g,n} - (\phi_g - \alpha_g)) \\
&\quad + \sum_{\text{group}(n)=2} y_{g,n}(\phi_g + \delta_g) - \sum_{\text{group}(n)=2} \exp(\rho_n + \varepsilon_{g,n} - (\phi_g + \delta_g)) \\
&\quad + \sum_{\text{group}(n)=3} y_{g,n}(\phi_g + \alpha_g) - \sum_{\text{group}(n)=3} \exp(\rho_n + \varepsilon_{g,n} - (\phi_g + \alpha_g)) \\
&\quad - \frac{(\phi_g - \theta_\phi)^2}{2\sigma_\phi^2} \\
&= \phi_g \sum_{\text{group}(n)=1} y_{g,n} - \alpha_g \sum_{\text{group}(n)=1} y_{g,n} - \exp(-\phi_g) \sum_{\text{group}(n)=1} \exp(\rho_n + \varepsilon_{g,n} + \alpha_g) \\
&\quad + \phi_g \sum_{\text{group}(n)=2} y_{g,n} + \delta_g \sum_{\text{group}(n)=2} y_{g,n} - \exp(-\phi_g) \sum_{\text{group}(n)=2} \exp(\rho_n + \varepsilon_{g,n} - \delta_g) \\
&\quad + \phi_g \sum_{\text{group}(n)=3} y_{g,n} + \alpha_g \sum_{\text{group}(n)=3} y_{g,n} - \exp(-\phi_g) \sum_{\text{group}(n)=3} \exp(\rho_n + \varepsilon_{g,n} - \alpha_g) \\
&\quad - \frac{(\phi_g - \theta_\phi)^2}{2\sigma_\phi^2} \\
&= N\overline{y}_{g.}\phi_g + S_2 - S\exp(-\phi_g) - \frac{(\phi_g - \theta_\phi)^2}{2\sigma_\phi^2}
\end{aligned}
$$

where

$$S_2 = \delta_g \sum_{\text{group}(n)=2} y_{g,n} + \alpha_g \left(\sum_{\text{group}(n)=3} y_{g,n} - \sum_{\text{group}(n)=1} y_{g,n} \right)$$
$$S = \sum_{\text{group}(n)=1} \exp(\rho_n + \varepsilon_{g,n} + \alpha_g) + \sum_{\text{group}(n)=2} \exp(\rho_n + \varepsilon_{g,n} - \delta_g) + \sum_{\text{group}(n)=3} \exp(\rho_n + \varepsilon_{g,n} - \alpha_g)$$

Differentiating g, we get

$$g'(\phi_g) = N\overline{y}_{g.} + S\exp(-\phi_g) - \frac{\phi_g - \theta_\phi}{\sigma_\phi}$$

We take $\widehat{\phi}_g$ to be the root of this derivative:

$$0 = N\overline{y}_{g.} + S\exp(-\widehat{\phi}_g) - \frac{\widehat{\phi}_g - \theta_\phi}{\sigma_\phi}$$

Taking the quadratic Taylor approximation of the exponential function,

$$\begin{aligned} 0 &= N\overline{y}_{g.} + S\left(1 - \widehat{\phi}_g - \frac{\widehat{\phi}_g^2}{2}\right) - \frac{\widehat{\phi}_g - \theta_\phi}{\sigma_\phi} \\ &= N\overline{y}_{g.} + S + \frac{\theta_\phi}{\sigma_\phi} + \left(-1 - \frac{1}{\sigma_\phi}\right)\widehat{\phi}_g + \left(-\frac{S}{2}\right)\widehat{\phi}_g^2 \end{aligned}$$

From the quadratic formula,

$$\widehat{\phi}_g = \frac{\left(1 + \frac{1}{\sigma_\phi}\right) \pm \sqrt{\left(1 + \frac{1}{\sigma_\phi}\right)^2 + 2S\left(N\overline{y}_{g.} + S + \frac{\theta_\phi}{\sigma_\phi}\right)}}{-S}$$

In practice, I will use the root with the higher value of $g(\widehat{\phi}_g)$.

C Old work: derivations of Metropolis proposals for point mass mixtures

C.1 α_g

I choose a proposal for α_g with the form,

$$q(\alpha_g \mid \theta'_\alpha, \sigma'_\alpha, \pi'_\alpha) = I(\alpha_g = 0)\pi'_\alpha + I(\alpha_g \neq 0)(1 - \pi'_\alpha)N(\alpha_g \mid \theta'_\alpha, (\sigma'_\alpha)^2),$$

which resembles the prior for α_g except that the parameters are updated to reflect the data, $\underline{y} = (y_{1,1}, \ldots, y_{G,N})$ (except for π'_α, for which we simply use π_α). To find θ'_α and σ'_α, we pretend that α_g has a $N(\alpha_g \mid \theta_\alpha, \sigma^2_\alpha)$ conditional likelihood, θ_α has a $N(\theta_\alpha \mid 0, c^2_\alpha)$ prior, and σ_α is fixed. From the rule on pages 46 and 47 of Gelman's book, the conditional posterior distribution of θ_α is

$$N\left(\theta_\alpha \left| \frac{\sigma_\alpha^{-2}\alpha_g}{c_\alpha^{-2} + \sigma_\alpha^{-2}} \ , \ (c_\alpha^{-2} + \sigma_\alpha^{-2})^{-1} \right.\right)$$

Hence, we let

$$\begin{aligned}
\theta'_\alpha &= \frac{\sigma_\alpha^{-2}\alpha_g}{c_\alpha^{-2} + \sigma_\alpha^{-2}} \\
(\sigma^2_\alpha)' &= \mathrm{Var}(\alpha_g) \\
&= \mathrm{Var}(E(\alpha_g \mid \theta_\alpha)) + E(\mathrm{Var}(\alpha_g \mid \theta_\alpha)) \\
&= \underbrace{\mathrm{Var}(\theta_\alpha)}_{\text{Use prior variance.}} + E(\sigma^2_\alpha) \\
&= c^2_\alpha + \sigma^2_\alpha
\end{aligned}$$

For example, whereas we interpret π_α as $P(\alpha = 0)$, a prior probability, we interpret π'_α as:

$$\begin{aligned}
\pi'_\alpha &= P(\alpha_g = 0 \mid \underline{y}, \ \ldots) \\
&= \frac{P(\underline{y} \mid \alpha_g = 0, \ldots)P(\alpha_g = 0)}{P(\underline{y} \mid \alpha_g = 0, \ \ldots)P(\alpha_g = 0) + P(\underline{y} \mid \alpha_g \neq 0, \ldots)P(\alpha_g \neq 0)} \\
&= \frac{1}{1 + \frac{P(\underline{y} \ \mid \ \alpha_g \neq 0, \ \ldots)}{P(\underline{y} \ \mid \ \alpha_g = 0, \ \ldots)} \frac{1-\pi_\alpha}{\pi_\alpha}} \\
&= \frac{1}{1 + \frac{1-\pi_\alpha}{\pi_\alpha} \prod_{k(n) \neq 2} \frac{P(y_{g,n} \ \mid \ \alpha_g \neq 0, \ \ldots)}{P(y_{g,n} \ \mid \ \alpha_g = 0, \ \ldots)}}
\end{aligned}$$

where "…" represents all the model parameters except for the other α_g's. To simplify the likelihood ratio in the denominator, we need $P(y_{g,n} \mid \alpha_g = 0, \ldots)$ and $P(y_{g,n} \mid \alpha_g \neq 0, \ldots)$.

$$\begin{aligned}
P(y_{g,n} \mid \alpha_g = 0, \ldots) &= \mathrm{Poisson}(y_{g,n} \mid \exp(\rho_n + \varepsilon_{g,n} + \mu(n, \phi_g, 0, \delta_g))) \\
&= \frac{1}{y_{g,n}!} \exp(-\exp(\rho_n + \varepsilon_{g,n} + \mu(n, \phi_g, 0, \delta_g))) \exp(y_{g,n} \cdot (\rho_n + \varepsilon_{g,n} + \mu(n, \phi_g, 0, \delta_g))) \\
&= \frac{1}{y_{g,n}!} \exp(y_{g,n} \cdot (\rho_n + \varepsilon_{g,n} + \mu(n, \phi_g, 0, \delta_g)) - \exp(\rho_n + \varepsilon_{g,n} + \mu(n, \phi_g, 0, \delta_g)))
\end{aligned}$$

I break up the calculation of $P(y_{g,n} \mid \alpha_g \neq 0, \ldots)$ into 2 cases.

1. Assume library n is in treatment group 1.

$$
\begin{aligned}
&P(y_{g,n} \mid \alpha_g \neq 0, \ldots) \\
&\quad = \int_{\alpha_g \neq 0} P(y_{g,n} \mid \alpha_g, \ldots) N(\alpha_g \mid \theta'_\alpha, (\sigma'_\alpha)^2) d\alpha_g \\
&\quad = \int_{\alpha_g \neq 0} \text{Poisson}(y_{g,n} \mid \exp(\rho_n + \varepsilon_{g,n} + \eta(g,n))) N(\alpha_g \mid \theta'_\alpha, (\sigma'_\alpha)^2) d\alpha_g \\
&\quad = \int_{\alpha_g \neq 0} \text{Poisson}(y_{g,n} \mid \exp(\underbrace{\rho_n + \varepsilon_{g,n} + \phi_g}_{i} - \alpha_g) N(\alpha_g \mid \theta'_\alpha, (\sigma'_\alpha)^2) d\alpha_g \\
&\quad = \int \frac{\exp(-\exp(i-\alpha_g))(\exp(i-\alpha_g))^{y_{g,n}}}{y_{g,n}!} \frac{1}{\sqrt{2\pi(\sigma'_\alpha)^2}} \exp\left(-\frac{(\alpha_g - \theta'_\alpha)^2}{2(\sigma_\alpha)'^2}\right) d\alpha_g \\
&\quad \approx \int \frac{\exp(-\frac{(i-\alpha_g)^2}{2} - (i-\alpha_g) - 1)(\exp(y_{g,n}(i-\alpha_g)))}{y_{g,n}!} \frac{1}{\sqrt{2\pi(\sigma'_\alpha)^2}} \exp\left(-\frac{(\alpha_g - \theta'_\alpha)^2}{2(\sigma_\alpha)'^2}\right) d\alpha_g \\
&\quad = (2\pi(\sigma'_\alpha)^2)^{-1/2}/y_{g,n}! \int \exp\left(-\frac{(i-\alpha_g)^2}{2} - i + \alpha_g - 1 + y_{g,n}(i-\alpha_g) - \frac{(\alpha_g - \theta'_\alpha)^2}{2(\sigma_\alpha)'^2}\right) d\alpha_g \\
&\quad = (2\pi(\sigma'_\alpha)^2)^{-1/2}/y_{g,n}! \int \exp\left(-\frac{\alpha_g^2}{2(\sigma'_\alpha)^2} - \frac{\alpha_g^2}{2} + i\alpha_g + \frac{\theta'_\alpha \alpha_g}{(\sigma'_\alpha)^2} - y_{g,n}\alpha_g + \alpha_g \right. \\
&\qquad \left. - \frac{i^2}{2} + iy_{g,n} - i - \frac{(\theta'_\alpha)^2}{2(\sigma'_\alpha)^2} - 1\right) d\alpha_g \\
&\quad = \underbrace{(2\pi(\sigma'_\alpha)^2)^{-1/2}/y_{g,n}!}_{D} \int \exp\left(\underbrace{\left(-\frac{1}{2(\sigma'_\alpha)^2} - \frac{1}{2}\right)}_{A} \alpha_g^2 + \underbrace{\left(i + \frac{\theta'_\alpha}{(\sigma'_\alpha)^2} - y_{g,n} + 1\right)}_{B} \alpha_g \right. \\
&\qquad \left. \underbrace{- \frac{i^2}{2} + iy_{g,n} - i - \frac{(\theta'_\alpha)^2}{2(\sigma'_\alpha)^2} - 1}_{C}\right) d\alpha_g \\
&\quad = D \int \exp(A\alpha_g^2 + B\alpha_g + C) d\alpha_g \\
&\quad = D \int \exp\left(A\left(\alpha_g + \frac{B}{2A}\right)^2 + C - \frac{B^2}{4A}\right) d\alpha_g \\
&\quad = D \exp\left(C - \frac{B^2}{4A}\right) \int \underbrace{\exp\left(-\frac{1}{2(1/(-2A))}\left(\alpha_g + \frac{B}{2A}\right)^2\right) d\alpha_g}_{\text{kernel of a normal distribution (note: } A<0\text{)}} \\
&\quad = D \exp\left(C - \frac{B^2}{4A}\right) \left(\frac{2\pi}{-2A}\right)^{1/2} \\
&\quad = D \exp\left(C - \frac{B^2}{4A}\right) \left(-\frac{\pi}{A}\right)^{1/2}
\end{aligned}
$$

2. $P(y_{g,n} \mid \alpha_g \neq 0, \ldots)$ is the same when n is in treatment group 3 except that B changes:

$$B = -i + \frac{\theta'_\alpha}{(\sigma'_\alpha)^2} + y_{g,n} - 1$$

C.2 δ_g

The proposal for δ_g is analogous to that of α_g:

$$q(\delta_g \mid \theta'_\delta, \sigma'_\delta, \pi'_\delta) = I(\delta_g = 0)\pi'_\delta + I(\delta_g \neq 0)(1 - \pi'_\delta)N(\delta_g \mid \theta'_\delta, (\sigma'_\delta)^2),$$

where:

$$\begin{aligned} \theta'_\delta &= \frac{\sigma_\delta^{-2}\delta_g}{c_\delta^{-2} + \sigma_\delta^{-2}} \\ (\sigma'_\delta)^2 &= c_\delta^2 + \sigma_\delta^2 \\ \pi'_\delta &= \pi_\delta \end{aligned}$$

$$\pi'_\delta = \frac{1}{1 + \frac{1-\pi_\delta}{\pi_\delta} \prod_{k(n)=2} \frac{P(y_{g,n} \ \mid \ \delta_g \neq 0, \ \ldots)}{P(y_{g,n} \ \mid \ \delta_g = 0, \ \ldots)}}$$

$$\begin{aligned} P(y_{g,n} \mid \delta_g = 0, \ldots) &= \frac{1}{y_{g,n}!} \exp(y_{g,n} \cdot (\rho_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, 0)) \\ &\quad - \exp(\rho_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, 0))) \\ P(y_{g,n} \mid \delta_g \neq 0, \ldots) &= D \exp\left(C - \frac{B^2}{4A}\right) \left(-\frac{\pi}{A}\right)^{1/2} \end{aligned}$$

$$\begin{aligned} A &= -\frac{1}{2(\sigma'_\delta)^2} - \frac{1}{2} \\ B &= -i + \frac{\theta'_\delta}{(\sigma'_\delta)^2} + y_{g,n} - 1 \\ C &= -\frac{i^2}{2} + i y_{g,n} - i - \frac{(\theta'_\delta)^2}{2(\sigma'_\delta)^2} - 1 \\ D &= (2\pi(\sigma'_\delta)^2)^{-1/2} / y_{g,n}! \end{aligned}$$

$$i = \rho_n + \varepsilon_{g,n} + \phi_g$$

$$\begin{aligned} \theta'_\delta &= \frac{c_\delta^{-2}\theta_\delta + \sigma_\delta^{-2} N_\delta^{-1} \sum_{k(n) \neq 2} y_{g,n}}{c_\delta^{-2} + \sigma_\delta^{-2}} \\ (\sigma'_\delta)^2 &= (c_\delta^{-2} + \sigma_\delta^{-2})^{-1} \end{aligned}$$

where N_δ is the number of libraries in the second treatment group.