

---

# A Fully Bayesian Model for RNA-seq Data

---

Will Landau

Department of Statistics  
Iowa State University

June 6, 2013

## Contents

<b>1</b>	<b>The Model</b>	<b>1</b>
<b>2</b>	<b>Full Conditional Distributions</b>	<b>3</b>
<b>A</b>	<b>Derivations of the Full Conditionals</b>	<b>4</b>
A.1	$p(c_n \mid \dots)$ : Metropolis . . . . .	5
A.2	$p\left(\frac{1}{\sigma_c^2} \mid \dots\right)$ : Truncated Gamma . . . . .	6
A.3	$p(\varepsilon_{g,n} \mid \dots)$ : Metropolis . . . . .	6
A.4	$p\left(\frac{1}{\sigma_g^2} \mid \dots\right)$ : Gamma . . . . .	7
A.5	$p(d \mid \dots)$ : Metropolis . . . . .	8
A.6	$p(\tau^2 \mid \dots)$ : Gamma . . . . .	8
A.7	$p(\phi_g \mid \dots)$ : Metropolis . . . . .	9
A.8	$p(\theta_\phi \mid \dots)$ : Normal . . . . .	9
A.9	$p\left(\frac{1}{\sigma_\phi^2} \mid \dots\right)$ : Truncated Gamma . . . . .	10
A.10	$p(\alpha_g \mid \dots)$ : Metropolis . . . . .	11
A.11	$p(\theta_\alpha \mid \dots)$ : Normal . . . . .	11
A.12	$p\left(\frac{1}{\sigma_\alpha} \mid \dots\right)$ : Truncated Gamma . . . . .	11
A.13	$p(\pi_\alpha \mid \dots)$ : Beta . . . . .	12
A.14	$p(\delta_g \mid \dots)$ : Metropolis . . . . .	12
A.15	$p(\theta_\delta \mid \dots)$ : Normal . . . . .	13
A.16	$p\left(\frac{1}{\sigma_\delta} \mid \dots\right)$ : Truncated Gamma . . . . .	13
A.17	$p(\pi_\delta \mid \dots)$ : Beta . . . . .	14

## 1 The Model

Let  $y_{g,n}$  be the expression level of gene  $g$  ( $g = 1, \dots, G$ ) in library  $n$  ( $n = 1, \dots, N$ ). Let  $\mu(n, \phi_g, \alpha_g, \delta_g)$  be the function given by:

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

Then:

$$\begin{aligned} y_{g,n} &\sim \text{Poisson}(\exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g))) \\ c_n &\sim \text{N}(0, \sigma_c^2) \\ \sigma_c &\sim \text{U}(0, \sigma_{c0}) \\ \varepsilon_{g,n} &\sim \text{N}(0, \sigma_g^2) \\ \sigma_g^2 &\sim \text{Inv-Gamma}\left(\text{shape} = \frac{d}{2}, \text{rate} = \frac{d \cdot \tau^2}{2}\right) \\ d &\sim \text{U}(0, d_0) \\ \tau^2 &\sim \text{Gamma}(\tau^2 \mid \text{shape} = a_\tau, \text{rate} = b_\tau) \\ \phi_g &\sim \text{N}(\theta_\phi, \sigma_\phi^2) \\ \theta_\phi &\sim \text{N}(0, \gamma_\phi^2) \\ \sigma_\phi &\sim \text{U}(0, \sigma_{\phi 0}) \\ \alpha_g &\sim \pi_\alpha^{1-I(\alpha_g)} [(1 - \pi_\alpha) \text{N}(\theta_\alpha, \sigma_\alpha^2)]^{I(\alpha_g)} \\ \theta_\alpha &\sim \text{N}(0, \gamma_\alpha^2) \\ \sigma_\alpha &\sim \text{U}(0, \sigma_{\alpha 0}) \\ \pi_\alpha &\sim \text{Beta}(a_\alpha, b_\alpha) \\ \delta_g &\sim \pi_\delta^{1-I(\delta_g)} [(1 - \pi_\delta) \text{N}(\theta_\delta, \sigma_\delta^2)]^{I(\delta_g)} \\ \theta_\delta &\sim \text{N}(0, \gamma_\delta^2) \\ \sigma_\delta &\sim \text{U}(0, \sigma_{\delta 0}) \\ \pi_\delta &\sim \text{Beta}(a_\delta, b_\delta) \end{aligned}$$

where:

- $I(x) = 0$  if  $x = 0$  and 1 otherwise.
- Independence is implied unless otherwise specified.
- The parameters to the left of the “ $\sim$ ” are implicitly conditioned on the parameters to the right.



## 2 Full Conditional Distributions

$$\begin{aligned}
p(c_n \mid \dots) &\propto \exp \left( \sum_{g=1}^G [y_{g,n} \log \lambda_{g,n} - \lambda_{g,n}] - \frac{c_n^2}{2\sigma_c^2} \right) \\
p \left( \frac{1}{\sigma_c^2} \mid \dots \right) &= \text{Gamma} \left( \frac{1}{\sigma_c^2} \mid \text{shape} = \frac{N-1}{2}, \text{rate} = \frac{1}{2} \sum_{n=1}^N c_n^2 \right) I \left( \frac{1}{\sigma_c^2} > \frac{1}{\sigma_{c0}^2} \right) \\
p(\varepsilon_{g,n} \mid \dots) &\propto \exp \left( y_{g,n} \log \lambda_{g,n} - \lambda_{g,n} - \frac{\varepsilon_{g,n}^2}{2\sigma_g^2} \right) \\
p \left( \frac{1}{\sigma_g^2} \mid \dots \right) &= \text{Gamma} \left( \frac{1}{\sigma_g^2} \mid \text{shape} = \frac{N+d}{2}, \text{rate} = \frac{1}{2} \left( d \cdot \tau^2 + \sum_{n=1}^N \varepsilon_{g,n}^2 \right) \right) \\
p(d \mid \dots) &\propto \Gamma(d/2)^{-G} \left( \frac{d \cdot \tau^2}{2} \right)^{Gd/2} \left( \prod_{g=1}^G \sigma_g^2 \right)^{-(d/2+1)} \exp \left( -\frac{d \cdot \tau^2}{2} \sum_{g=1}^G \frac{1}{\sigma_g^2} \right) I(0 < d < d_0) \\
p(\tau^2 \mid \dots) &= \text{Gamma} \left( \tau^2 \mid \text{shape} = a_\tau + \frac{Gd}{2}, \text{rate} = b_\tau + \frac{d}{2} \sum_{g=1}^G \frac{1}{\sigma_g^2} \right) \\
p(\phi_g \mid \dots) &\propto \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) \\
p(\theta_\phi \mid \dots) &= N \left( \frac{\gamma_\phi^2 \sum_{g=1}^G \phi_g}{G\gamma_\phi^2 + \sigma_\phi^2}, \frac{\gamma_\phi^2 \sigma_\phi^2}{G\gamma_\phi^2 + \sigma_\phi^2} \right) \\
p \left( \frac{1}{\sigma_\phi^2} \mid \dots \right) &= \text{Gamma} \left( \text{shape} = \frac{G-1}{2}, \text{rate} = \frac{1}{2} \sum_{g=1}^G (\phi_g - \theta_\phi)^2 \right) I \left( \frac{1}{\sigma_\phi^2} > \frac{1}{\sigma_{\phi0}^2} \right) \\
p(\alpha_g \mid \dots) &\propto \exp \left( \sum_{k(n) \neq 2} [y_{g,n} \log \lambda_{g,n} - \lambda_{g,n}] - I(\alpha_g \neq 0) \frac{(\alpha_g - \theta_\alpha)^2}{2\sigma_\alpha^2} \right) \\
p(\theta_\alpha \mid \dots) &= N \left( \frac{\gamma_\alpha^2 \sum_{\alpha_g \neq 0} \alpha_g}{G_\alpha \gamma_\alpha^2 + \sigma_\alpha^2}, \frac{\gamma_\alpha^2 \sigma_\alpha^2}{G_\alpha \gamma_\alpha^2 + \sigma_\alpha^2} \right) \\
p \left( \frac{1}{\sigma_\alpha^2} \mid \dots \right) &= \text{Gamma} \left( \frac{1}{\sigma_\alpha^2} \mid \text{shape} = \frac{G_\alpha - 1}{2}, \text{rate} = \frac{1}{2} \sum_{\alpha_g \neq 0} (\alpha_g - \theta_\alpha)^2 \right) \\
p(\pi_\alpha \mid \dots) &= \text{Beta}(G - G_\alpha + \alpha_\tau, G_\alpha + b_\tau) \\
p(\delta_g \mid \dots) &\propto \exp \left( \sum_{k(n)=2} [y_{g,n} \log \lambda_{g,n} - \lambda_{g,n}] - I(\delta_g \neq 0) \frac{(\delta_g - \theta_\delta)^2}{2\sigma_\delta^2} \right) \\
p(\theta_\delta \mid \dots) &= N \left( \frac{\gamma_\delta^2 \sum_{\delta_g \neq 0} \delta_g}{G_\delta \gamma_\delta^2 + \sigma_\delta^2}, \frac{\gamma_\delta^2 \sigma_\delta^2}{G_\delta \gamma_\delta^2 + \sigma_\delta^2} \right) \\
p \left( \frac{1}{\sigma_\delta^2} \mid \dots \right) &= \text{Gamma} \left( \frac{1}{\sigma_\delta^2} \mid \text{shape} = \frac{G_\delta - 1}{2}, \text{rate} = \frac{1}{2} \sum_{\delta_g \neq 0} (\delta_g - \theta_\delta)^2 \right) \\
p(\pi_\delta \mid \dots) &= \text{Beta}(G - G_\delta + \delta_\tau, G_\delta + b_\tau)
\end{aligned}$$

## A Derivations of the Full Conditionals

Let  $k(n)$  be the treatment group of library  $n$ . Then from the model in Section 1, we get:

$$\begin{aligned}
p(c_n \mid \dots) &= \left[ \prod_{g=1}^G \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g))) \right] \cdot \text{N}(c_n \mid 0, \sigma_c^2) \\
p(\sigma_c \mid \dots) &= \left[ \prod_{n=1}^N \text{N}(c_n \mid 0, \sigma_c^2) \right] \cdot \text{U}(\sigma_c \mid 0, \sigma_{c0}) \\
p(\varepsilon_{g,n} \mid \dots) &= \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g))) \cdot \text{N}(\varepsilon_{g,n} \mid 0, \sigma_g^2) \\
p(\sigma_g^2 \mid \dots) &= \left[ \prod_{n=1}^N \text{N}(\varepsilon_{g,n} \mid 0, \sigma_g^2) \right] \cdot \text{Inv-Gamma} \left( \sigma_g^2 \mid \text{shape} = \frac{d}{2}, \text{rate} = \frac{d \cdot \tau^2}{2} \right) \\
p(d \mid \dots) &= \left[ \prod_{g=1}^G \text{Inv-Gamma} \left( \sigma_g^2 \mid \text{shape} = \frac{d}{2}, \text{rate} = \frac{d \cdot \tau^2}{2} \right) \right] \cdot \text{U}(d \mid 0, d_0) \\
p(\tau^2 \mid \dots) &= \left[ \prod_{g=1}^G \text{Inv-Gamma} \left( \sigma_g^2 \mid \text{shape} = \frac{d}{2}, \text{rate} = \frac{d \cdot \tau^2}{2} \right) \right] \cdot \text{Gamma}(\tau^2 \mid \text{shape} = a_\tau, \text{rate} = b_\tau) \\
p(\phi_g \mid \dots) &= \left[ \prod_{n=1}^N \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g))) \right] \cdot \text{N}(\phi_g \mid \theta_\phi, \sigma_\phi^2) \\
p(\theta_\phi \mid \dots) &= \left[ \prod_{g=1}^G \text{N}(\phi_g \mid \theta_\phi, \sigma_\phi^2) \right] \cdot \text{N}(\theta_\phi \mid 0, \gamma_\phi^2) \\
p(\sigma_\phi \mid \dots) &= \left[ \prod_{g=1}^G \text{N}(\phi_g \mid \theta_\phi, \sigma_\phi^2) \right] \cdot \text{U}(\sigma_\phi \mid 0, \sigma_{\phi 0})
\end{aligned}$$

$$\begin{aligned}
p(\alpha_g \mid \dots) &= \left[ \prod_{k(n) \neq 2} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g))) \right] \pi_\alpha^{1-I(\alpha_g)} [(1 - \pi_\alpha) \text{N}(\alpha_g \mid \theta_\alpha, \sigma_\alpha^2)]^{I(\alpha_g)} \\
p(\theta_\alpha \mid \dots) &= \left[ \prod_{g=1}^G \pi_\alpha^{1-I(\alpha_g)} [(1 - \pi_\alpha) \text{N}(\alpha_g \mid \theta_\alpha, \sigma_\alpha^2)]^{I(\alpha_g)} \right] \cdot \text{N}(\theta_\alpha \mid 0, \gamma_\alpha^2) \\
p(\sigma_\alpha \mid \dots) &= \left[ \prod_{g=1}^G \pi_\alpha^{1-I(\alpha_g)} [(1 - \pi_\alpha) \text{N}(\alpha_g \mid \theta_\alpha, \sigma_\alpha^2)]^{I(\alpha_g)} \right] \cdot \text{U}(\sigma_\alpha \mid 0, \sigma_{\alpha 0}) \\
p(\pi_\alpha \mid \dots) &= \left[ \prod_{g=1}^G \pi_\alpha^{1-I(\alpha_g)} [(1 - \pi_\alpha) \text{N}(\alpha_g \mid \theta_\alpha, \sigma_\alpha^2)]^{I(\alpha_g)} \right] \cdot \text{Beta}(\pi_\alpha \mid a_\alpha, b_\alpha) \\
p(\delta_g \mid \dots) &= \left[ \prod_{k(n)=2} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \delta_g, \delta_g))) \right] \pi_\delta^{1-I(\delta_g)} [(1 - \pi_\delta) \text{N}(\delta_g \mid \theta_\delta, \sigma_\delta^2)]^{I(\delta_g)} \\
p(\theta_\delta \mid \dots) &= \left[ \prod_{g=1}^G \pi_\delta^{1-I(\delta_g)} [(1 - \pi_\delta) \text{N}(\delta_g \mid \theta_\delta, \sigma_\delta^2)]^{I(\delta_g)} \right] \cdot \text{N}(\theta_\delta \mid 0, \gamma_\delta^2) \\
p(\sigma_\delta \mid \dots) &= \left[ \prod_{g=1}^G \pi_\delta^{1-I(\delta_g)} [(1 - \pi_\delta) \text{N}(\delta_g \mid \theta_\delta, \sigma_\delta^2)]^{I(\delta_g)} \right] \cdot \text{U}(\sigma_\delta \mid 0, \sigma_{\delta 0}) \\
p(\pi_\delta \mid \dots) &= \left[ \prod_{g=1}^G \pi_\delta^{1-I(\delta_g)} [(1 - \pi_\delta) \text{N}(\delta_g \mid \theta_\delta, \sigma_\delta^2)]^{I(\delta_g)} \right] \cdot \text{Beta}(\pi_\delta \mid a_\delta, b_\delta)
\end{aligned}$$

### A.1 $p(c_n \mid \dots)$ : Metropolis

$$p(c_n \mid \dots) = \left[ \prod_{g=1}^G \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g))) \right] \cdot \text{N}(c_n \mid 0, \sigma_c^2)$$

which is nothing immediately recognizable. I will sample from this distribution in a Metropolis step using the kernel of this distribution (taking  $\lambda_{g,n} = \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g))$ ):

$$\begin{aligned}
&\left[ \prod_{g=1}^G \lambda_{g,n}^{y_{g,n}} \exp(-\lambda_{g,n}) \right] \exp\left(-\frac{c_n^2}{2\sigma_c^2}\right) \\
&= \exp\left(\sum_{g=1}^G [y_{g,n} \log \lambda_{g,n} - \lambda_{g,n}] - \frac{c_n^2}{2\sigma_c^2}\right)
\end{aligned}$$

where the sum inside the exponent can be parallelized on the GPU.

## A.2 $p\left(\frac{1}{\sigma_c^2} \mid \dots\right)$ Truncated Gamma

$$\begin{aligned} p(\sigma_c \mid \dots) &= \left[ \prod_{n=1}^N N(c_n \mid 0, \sigma_c^2) \right] \cdot U(\sigma_c \mid 0, \sigma_{c0}) \\ &\propto \prod_{n=1}^N \frac{1}{\sqrt{\sigma_c^2}} \exp\left(-\frac{c_n^2}{2\sigma_c^2}\right) \cdot I(0 < \sigma_c < \sigma_{c0}) \\ &= (\sigma_c^2)^{-N/2} \exp\left(-\frac{1}{\sigma_c^2} \frac{1}{2} \sum_{n=1}^N c_n^2\right) \cdot I(0 < \sigma_c < \sigma_{c0}) \end{aligned}$$

which, for constants  $a$  and  $b = \frac{1}{2} \sum_{n=1}^N c_n^2$ , can be written as

$$p(\sigma_c \mid \dots) = a \cdot (\sigma_c^2)^{-N/2} \exp\left(-\frac{1}{\sigma_c^2} b\right) I(0 < \sigma_c^2 < \sigma_{c0}^2)$$

Transformation: let  $z = g(\sigma_c) = \sigma_c^2$  so that  $g^{-1}(z) = \sqrt{z}$  and:

$$\begin{aligned} p(\sigma_c^2 = z \mid \dots) &= p(\sigma_c = g^{-1}(z) \mid \dots) \left| \frac{dg^{-1}(z)}{dz} \right| \\ &= a \cdot z^{-N/2} \exp\left(-\frac{1}{(\sqrt{z})^2} b\right) I(0 < z < \sigma_{c0}^2) \left| -\frac{1}{2} z^{-1/2} \right| \\ &= \frac{a}{2} z^{-(N/2-1/2+1)} \exp\left(-\frac{1}{z} b\right) I(0 < z < \sigma_{c0}^2) \\ &= \text{Inv-Gamma}\left(z \mid \text{shape} = \frac{N-1}{2}, \text{rate} = b\right) I(0 < z < \sigma_{c0}^2) \end{aligned}$$

Recalling that  $b = \frac{1}{2} \sum_{n=1}^N c_n^2$ ,

$$p\left(\frac{1}{\sigma_c^2} \mid \dots\right) = \text{Gamma}\left(\frac{1}{\sigma_c^2} \mid \text{shape} = \frac{N-1}{2}, \text{rate} = \frac{1}{2} \sum_{n=1}^N c_n^2\right) I\left(\frac{1}{\sigma_c^2} > \frac{1}{\sigma_{c0}^2}\right)$$

## A.3 $p(\varepsilon_{g,n} \mid \dots)$ : Metropolis

$$p(\varepsilon_{g,n} \mid \dots) = \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g))) \cdot N(\varepsilon_{g,n} \mid 0, \sigma_g^2)$$



which is nothing immediately recognizable. I will sample from this distribution in a Metropolis step using the kernel:

$$\begin{aligned} & \lambda_{g,n}^{y_{g,n}} \exp(-\lambda_{g,n}) \exp\left(-\frac{\varepsilon_{g,n}^2}{2\sigma_g^2}\right) \\ &= \exp\left(y_{g,n} \log \lambda_{g,n} - \lambda_{g,n} - \frac{\varepsilon_{g,n}^2}{2\sigma_g^2}\right) \end{aligned}$$

where  $\lambda_{g,n} = \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g))$ . The  $\varepsilon_{g,n}$ 's will be sampled in parallel across genes on the GPU.

#### A.4 $p\left(\frac{1}{\sigma_g^2} \mid \dots\right)$ Gamma

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

The last line is the kernel of an inverse gamma distribution with shape parameter  $\frac{N+d}{2}$  and rate parameter  $\frac{1}{2} \left(d \cdot \tau^2 + \sum_{n=1}^N \varepsilon_{g,n}^2\right)$ . Hence:

$$p\left(\frac{1}{\sigma_g^2} \mid \dots\right) = \text{Gamma}\left(\frac{1}{\sigma_g^2} \mid \text{shape} = \frac{N+d}{2}, \text{rate} = \frac{1}{2} \left(d \cdot \tau^2 + \sum_{n=1}^N \varepsilon_{g,n}^2\right)\right)$$

The  $1/\sigma_g^2$ 's will be sampled in parallel on the GPU.

### A.5 $p(d \mid \dots)$ : Metropolis

$$\begin{aligned}
p(d \mid \dots) &= \left[ \prod_{g=1}^G \text{Inv-Gamma} \left( \sigma_g^2 \mid \text{shape} = \frac{d}{2}, \text{rate} = \frac{d \cdot \tau^2}{2} \right) \right] \cdot \text{U}(d \mid 0, d_0) \\
&\propto \prod_{g=1}^G \Gamma(d/2)^{-1} \left( \frac{d \cdot \tau^2}{2} \right)^{d/2} (\sigma_g^2)^{-(d/2+1)} \exp \left( -\frac{1}{\sigma_g^2} \frac{d \cdot \tau^2}{2} \right) I(0 < d < d_0) \\
&\propto \Gamma(d/2)^{-G} \left( \frac{d \cdot \tau^2}{2} \right)^{Gd/2} \left( \prod_{g=1}^G \sigma_g^2 \right)^{-(d/2+1)} \exp \left( -\frac{d \cdot \tau^2}{2} \sum_{g=1}^G \frac{1}{\sigma_g^2} \right) I(0 < d < d_0)
\end{aligned}$$

I will sample  $d$  with a Metropolis step using the above kernel. Sums and products over  $g$  ( $g = 1, \dots, G$ ) will be done in parallel on the GPU.

### A.6 $p(\tau^2 \mid \dots)$ : Gamma

$$\begin{aligned}
p(\tau^2 \mid \dots) &= \left[ \prod_{g=1}^G \text{Inv-Gamma} \left( \sigma_g^2 \mid \text{shape} = \frac{d}{2}, \text{rate} = \frac{d \cdot \tau^2}{2} \right) \right] \cdot \text{Gamma}(\tau^2 \mid \text{shape} = a_\tau, \text{rate} = b_\tau) \\
&\propto \left[ \Gamma(d/2)^{-G} \left( \frac{d \cdot \tau^2}{2} \right)^{Gd/2} \left( \prod_{g=1}^G \sigma_g^2 \right)^{-(d/2+1)} \exp \left( -\frac{d \cdot \tau^2}{2} \sum_{g=1}^G \frac{1}{\sigma_g^2} \right) \right] \cdot (\tau^2)^{a_\tau-1} \exp(-b_\tau \tau^2) \\
&\propto \left[ (\tau^2)^{Gd/2} \exp \left( -\tau^2 \cdot \frac{d}{2} \sum_{g=1}^G \frac{1}{\sigma_g^2} \right) \right] \cdot (\tau^2)^{a_\tau-1} \exp(-b_\tau \tau^2) \\
&= (\tau^2)^{Gd/2+a_\tau-1} \exp \left( -\tau^2 \left( b_\tau + \frac{d}{2} \sum_{g=1}^G \frac{1}{\sigma_g^2} \right) \right)
\end{aligned}$$

Hence:

$$p(\tau^2 \mid \dots) = \text{Gamma} \left( \tau^2 \mid \text{shape} = a_\tau + \frac{Gd}{2}, \text{rate} = b_\tau + \frac{d}{2} \sum_{g=1}^G \frac{1}{\sigma_g^2} \right)$$

### A.7 $p(\phi_g \mid \dots)$ : Metropolis

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

where  $\lambda_{g,n} = \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g))$ . I will sample the  $\phi_g$ 's in parallel using Metropolis steps.

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

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

Hence:

$$p(\theta_\phi \mid \dots) = N\left(\frac{\gamma_\phi^2 \sum_{g=1}^G \phi_g}{G\gamma_\phi^2 + \sigma_\phi^2}, \frac{\gamma_\phi^2 \sigma_\phi^2}{G\gamma_\phi^2 + \sigma_\phi^2}\right)$$

### A.9 $p\left(\frac{1}{\sigma_\phi^2} \mid \dots\right)$ : Truncated Gamma

$$\begin{aligned} p(\sigma_\phi \mid \dots) &= \left[ \prod_{g=1}^G N(\phi_g \mid \theta_\phi, \sigma_\phi^2) \right] \cdot U(\sigma_\phi \mid 0, \sigma_{\phi 0}) \\ &\propto \prod_{g=1}^G (\sigma_\phi^2)^{-1/2} \exp\left(-\frac{(\phi_g - \theta_\phi)^2}{2\sigma_\phi^2}\right) I(0 < \sigma_\phi^2 < \sigma_{\phi 0}^2) \\ &= (\sigma_\phi^2)^{-G/2} \exp\left(-\sum_{g=1}^G \frac{(\phi_g - \theta_\phi)^2}{2\sigma_\phi^2}\right) I(0 < \sigma_\phi^2 < \sigma_{\phi 0}^2) \\ &= (\sigma_\phi^2)^{-G/2} \exp\left(-\frac{1}{\sigma_\phi^2} \frac{1}{2} \sum_{g=1}^G (\phi_g - \theta_\phi)^2\right) I(0 < \sigma_\phi^2 < \sigma_{\phi 0}^2) \end{aligned}$$

Transformation: let  $z = g(\sigma_\phi) = \sigma_\phi^2$  so that  $g^{-1}(z) = \sqrt{z}$ . Then for some proportionality constant,  $a$ :

$$\begin{aligned} p(\sigma_\phi^2 = z \mid \dots) &= p(\sigma_\phi = g^{-1}(z) \mid \dots) \left| \frac{g^{-1}(z)}{dz} \right| \\ &= p(\sigma_\phi = \sqrt{z} \mid \dots) \left| \frac{1}{2} z^{-1/2} \right| \\ &= a \cdot (z)^{-G/2} \exp\left(-\frac{1}{z} \frac{1}{2} \sum_{g=1}^G (\phi_g - \theta_\phi)^2\right) I(0 < z < \sigma_{\phi 0}^2) z^{-1/2} \\ &= \frac{a}{2} \cdot z^{-(G/2-1/2+1)} \exp\left(-\frac{1}{z} \frac{1}{2} \sum_{g=1}^G (\phi_g - \theta_\phi)^2\right) I(0 < z < \sigma_{\phi 0}^2) \end{aligned}$$

which is a truncated inverse gamma distribution with shape parameter  $\frac{G-1}{2}$  and rate parameter  $\frac{1}{2} \sum_{g=1}^G (\phi_g - \theta_\phi)^2$ . Thus:

$$p\left(\frac{1}{\sigma_\phi^2} \mid \dots\right) = \text{Gamma}\left(\text{shape} = \frac{G-1}{2}, \text{rate} = \frac{1}{2} \sum_{g=1}^G (\phi_g - \theta_\phi)^2\right) I\left(\frac{1}{\sigma_\phi^2} > \frac{1}{\sigma_{\phi 0}^2}\right)$$

I will sample  $1/\sigma_\phi^2$  using the inverse cdf method.

### A.10 $p(\alpha_g \mid \dots)$ : Metropolis

$$\begin{aligned}
p(\alpha_g \mid \dots) &= \left[ \prod_{k(n) \neq 2} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g))) \right] \pi_\alpha^{1-I(\alpha_g)} [(1 - \pi_\alpha) \text{N}(\theta_\alpha, \sigma_\alpha^2)]^{I(\alpha_g)} \\
&\propto \left[ \prod_{k(n) \neq 2} \lambda_{g,n}^{y_{g,n}} \exp(-\lambda_{g,n}) \right] \exp \left( -I(\alpha_g) \frac{(\alpha_g - \theta_\alpha)^2}{2\sigma_\alpha^2} \right) \\
&= \exp \left( \sum_{k(n) \neq 2} [y_{g,n} \log \lambda_{g,n} - \lambda_{g,n}] - I(\alpha_g) \frac{(\alpha_g - \theta_\alpha)^2}{2\sigma_\alpha^2} \right)
\end{aligned}$$

### A.11 $p(\theta_\alpha \mid \dots)$ : Normal

$$\begin{aligned}
p(\theta_\alpha \mid \dots) &= \left[ \prod_{g=1}^G \pi_\alpha^{1-I(\alpha_g)} [(1 - \pi_\alpha) \text{N}(\alpha_g \mid \theta_\alpha, \sigma_\alpha^2)]^{I(\alpha_g)} \right] \cdot \text{N}(\theta_\alpha \mid 0, \gamma_\alpha^2) \\
&\propto \left[ \prod_{\alpha_g \neq 0} \text{N}(\alpha_g \mid \theta_\alpha, \sigma_\alpha^2) \right] \cdot \text{N}(\theta_\alpha \mid 0, \gamma_\alpha^2)
\end{aligned}$$

From algebra similar to the derivation of  $p(\theta_\phi \mid \dots)$ , we get:

$$p(\theta_\alpha \mid \dots) = N \left( \frac{\gamma_\alpha^2 \sum_{\alpha_g \neq 0} \alpha_g}{G_\alpha \gamma_\alpha^2 + \sigma_\alpha^2}, \frac{\gamma_\alpha^2 \sigma_\alpha^2}{G_\alpha \gamma_\alpha^2 + \sigma_\alpha^2} \right)$$

where  $G_\alpha$  is the number of genes for which  $\alpha_g \neq 0$ .

### A.12 $p(\frac{1}{\sigma_\alpha} \mid \dots)$ : Truncated Gamma

$$\begin{aligned}
p(\sigma_\alpha \mid \dots) &= \left[ \prod_{g=1}^G \pi_\alpha^{1-I(\alpha_g)} [(1 - \pi_\alpha) \text{N}(\alpha_g \mid \theta_\alpha, \sigma_\alpha^2)]^{I(\alpha_g)} \right] \cdot \text{U}(\sigma_\alpha \mid 0, \sigma_{\alpha 0}) \\
&\propto \prod_{\alpha_g \neq 0} \text{N}(\alpha_g \mid \theta_\alpha, \sigma_\alpha^2) \cdot I(0 < \sigma_\alpha^2 < \sigma_{\alpha 0}^2) \\
&\propto \prod_{\alpha_g \neq 0} (\sigma_\alpha^2)^{-1/2} \exp \left( -\frac{(\alpha_g - \theta_\alpha)^2}{2\sigma_\alpha^2} \right) \cdot I(0 < \sigma_\alpha^2 < \sigma_{\alpha 0}^2) \\
&= (\sigma_\alpha^2)^{-G_\alpha/2} \exp \left( -\frac{1}{\theta_\alpha^2} \frac{1}{2} \sum_{\alpha_g \neq 0} (\alpha_g - \theta_\alpha)^2 \right) \cdot I(0 < \sigma_\alpha^2 < \sigma_{\alpha 0}^2)
\end{aligned}$$

where  $G_\alpha$  is the number of genes for which  $\alpha_g \neq 0$ . Transformation: let  $z = \sigma_\alpha^2 = g(\sigma_\alpha)$  so that  $g^{-1}(z) = \sqrt{z}$ . Then:

$$\begin{aligned} p(\sigma_\alpha^2 = z \mid \dots) &= p(\sigma_\alpha = \sqrt{z} \mid \dots) \left| \frac{d}{dz} \sqrt{z} \right|_z \\ &= (z)^{-G_\alpha/2} \exp \left( -\frac{1}{z} \frac{1}{2} \sum_{\alpha_g \neq 0} (\alpha_g - \theta_\alpha)^2 \right) \cdot I(0 < z < \sigma_{\alpha 0}^2) \left| -\frac{1}{2} z^{-1/2} \right| \\ &= (z)^{-(G_\alpha/2 - 1/2 + 1)} \exp \left( -\frac{1}{z} \frac{1}{2} \sum_{\alpha_g \neq 0} (\alpha_g - \theta_\alpha)^2 \right) \cdot I(0 < z < \sigma_{\alpha 0}^2) \end{aligned}$$

which is the kernel of an Inverse-Gamma distribution. Hence:

$$p\left(\frac{1}{\sigma_\alpha^2} \mid \dots\right) = \text{Gamma} \left( \frac{1}{\sigma_\alpha^2} \mid \text{shape} = \frac{G_\alpha - 1}{2}, \text{rate} = \frac{1}{2} \sum_{\alpha_g \neq 0} (\alpha_g - \theta_\alpha)^2 \right)$$

#### A.13 $p(\pi_\alpha \mid \dots)$ : Beta

$$\begin{aligned} p(\pi_\alpha \mid \dots) &= \left[ \prod_{g=1}^G \pi_\alpha^{1-I(\alpha_g)} [(1 - \pi_\alpha) \text{N}(\alpha_g \mid \theta_\alpha, \sigma_\alpha^2)]^{I(\alpha_g)} \right] \cdot \text{Beta}(\pi_\alpha \mid a_\alpha, b_\alpha) \\ &\propto [\pi_\alpha^{G-G_\alpha} (1 - \pi_\alpha)^{G_\alpha}] \pi_\alpha^{a_\tau-1} (1 - \pi_\alpha)^{b_\tau-1} \\ &= \pi_\alpha^{G-G_\alpha+a_\tau-1} (1 - \pi_\alpha)^{G_\alpha+b_\tau-1} \end{aligned}$$

where  $G_\alpha$  is the number of genes for which  $\alpha_g \neq 0$ . Hence:

$$p(\pi_\alpha \mid \dots) = \text{Beta}(G - G_\alpha + \alpha_\tau, G_\alpha + b_\tau)$$

#### A.14 $p(\delta_g \mid \dots)$ : Metropolis

$$\begin{aligned} p(\delta_g \mid \dots) &= \left[ \prod_{k(n)=2} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \delta_g, \delta_g))) \right] \pi_\delta^{1-I(\delta_g)} [(1 - \pi_\delta) \text{N}(\theta_\delta, \sigma_\delta^2)]^{I(\delta_g)} \\ &\propto \left[ \prod_{k(n)=2} \lambda_{g,n}^{y_{g,n}} \exp(-\lambda_{g,n}) \right] \exp \left( -I(\delta_g) \frac{(\delta_g - \theta_\delta)^2}{2\sigma_\delta^2} \right) \\ &= \exp \left( \sum_{k(n)=2} [y_{g,n} \log \lambda_{g,n} - \lambda_{g,n}] - I(\delta_g) \frac{(\delta_g - \theta_\delta)^2}{2\sigma_\delta^2} \right) \end{aligned}$$

### A.15 $p(\theta_\delta \mid \dots)$ : Normal

$$\begin{aligned} p(\theta_\delta \mid \dots) &= \left[ \prod_{g=1}^G \pi_\delta^{1-I(\delta_g)} [(1 - \pi_\delta) \text{N}(\delta_g \mid \theta_\delta, \sigma_\delta^2)]^{I(\delta_g)} \right] \cdot \text{N}(\theta_\delta \mid 0, \gamma_\delta^2) \\ &\propto \left[ \prod_{\delta_g \neq 0} \text{N}(\delta_g \mid \theta_\delta, \sigma_\delta^2) \right] \cdot \text{N}(\theta_\delta \mid 0, \gamma_\delta^2) \end{aligned}$$

From algebra similar to the derivation of  $p(\theta_\phi \mid \dots)$ , we get:

$$p(\theta_\delta \mid \dots) = N \left( \frac{\gamma_\delta^2 \sum_{\delta_g \neq 0} \delta_g}{G_\delta \gamma_\delta^2 + \sigma_\delta^2}, \frac{\gamma_\delta^2 \sigma_\delta^2}{G_\delta \gamma_\delta^2 + \sigma_\delta^2} \right)$$

where  $G_\delta$  is the number of genes for which  $\delta_g \neq 0$ .

### A.16 $p(\frac{1}{\sigma_\delta} \mid \dots)$ : Truncated Gamma

$$\begin{aligned} p(\sigma_\delta \mid \dots) &= \left[ \prod_{g=1}^G \pi_\delta^{1-I(\delta_g)} [(1 - \pi_\delta) \text{N}(\delta_g \mid \theta_\delta, \sigma_\delta^2)]^{I(\delta_g)} \right] \cdot \text{U}(\sigma_\delta \mid 0, \sigma_{\delta 0}^2) \\ &\propto \prod_{\delta_g \neq 0} \text{N}(\delta_g \mid \theta_\delta, \sigma_\delta^2) \cdot I(0 < \sigma_\delta^2 < \sigma_{\delta 0}^2) \\ &\propto \prod_{\delta_g \neq 0} (\sigma_\delta^2)^{-1/2} \exp \left( -\frac{(\delta_g - \theta_\delta)^2}{2\sigma_\delta^2} \right) \cdot I(0 < \sigma_\delta^2 < \sigma_{\delta 0}^2) \\ &= (\sigma_\delta^2)^{-G_\delta/2} \exp \left( -\frac{1}{\theta_\delta^2} \frac{1}{2} \sum_{\delta_g \neq 0} (\delta_g - \theta_\delta)^2 \right) \cdot I(0 < \sigma_\delta^2 < \sigma_{\delta 0}^2) \end{aligned}$$

where  $G_\delta$  is the number of genes for which  $\delta_g \neq 0$ . Transformation: let  $z = \sigma_\delta^2 = g(\sigma_\delta)$  so that  $g^{-1}(z) = \sqrt{z}$ . Then:

$$\begin{aligned} p(\sigma_\delta^2 = z \mid \dots) &= p(\sigma_\delta = \sqrt{z} \mid \dots) \left| \frac{d}{dz} \sqrt{z} \right|_z \\ &= (z)^{-G_\delta/2} \exp \left( -\frac{1}{z} \frac{1}{2} \sum_{\delta_g \neq 0} (\delta_g - \theta_\delta)^2 \right) \cdot I(0 < z < \sigma_{\delta 0}^2) \left| -\frac{1}{2} z^{-1/2} \right| \\ &= (z)^{-(G_\delta/2 - 1/2 + 1)} \exp \left( -\frac{1}{z} \frac{1}{2} \sum_{\delta_g \neq 0} (\delta_g - \theta_\delta)^2 \right) \cdot I(0 < z < \sigma_{\delta 0}^2) \end{aligned}$$

which is the kernel of an Inverse-Gamma distribution. Hence:

$$p\left(\frac{1}{\sigma_\delta^2} \mid \cdots\right) = \text{Gamma}\left(\frac{1}{\sigma_\delta^2} \mid \text{shape} = \frac{G_\delta - 1}{2}, \text{rate} = \frac{1}{2} \sum_{\delta_g \neq 0} (\delta_g - \theta_\delta)^2\right)$$

**A.17**  $p(\pi_\delta \mid \cdots)$ : **Beta**

$$\begin{aligned} p(\pi_\delta \mid \cdots) &= \left[ \prod_{g=1}^G \pi_\delta^{1-I(\delta_g)} [(1 - \pi_\delta) \text{N}(\delta_g \mid \theta_\delta, \sigma_\delta^2)]^{I(\delta_g)} \right] \cdot \text{Beta}(\pi_\delta \mid a_\delta, b_\delta) \\ &\propto [\pi_\delta^{G-G_\delta} (1 - \pi_\delta)^{G_\delta}] \pi_\delta^{a_\tau-1} (1 - \pi_\delta)^{b_\tau-1} \\ &= \pi_\delta^{G-G_\delta+a_\tau-1} (1 - \pi_\delta)^{G_\delta+b_\tau-1} \end{aligned}$$

where  $G_\delta$  is the number of genes for which  $\delta_g \neq 0$ . Hence:

$$p(\pi_\delta \mid \cdots) = \text{Beta}(G - G_\delta + \delta_\tau, G_\delta + b_\tau)$$