between analytical mode and iteratively computed mode.

Algorithm 1: Dirichlet-Poisson-Binomial posterior mode estimation via the Lindley-Smith algorithm

//Initialization
Set the initial values $\gamma_1, \gamma_2, \mathbf{N}, \mathbf{m}_0, \mathbf{g}$.
//Main loop

while $\sum |m_i^{new} - m_i| > 10^{-5}$ **do**
  $m_i^{\text{new}} = \max[\pi(m_i|t_i, g_i, N, \boldsymbol{\phi}, \boldsymbol{\alpha}, \gamma_1, \gamma_2] = (g_i + \alpha)/(N)$
  $g_i^{\text{new}} = \max[\pi(g_i|\mathbf{T}, \mathbf{m}, N, \boldsymbol{\phi}, \boldsymbol{\alpha}, \gamma_1, \gamma_2] = t_i + \text{floor}(N m_i(1 - \phi_i))$
  $N^{\text{new}} = \max[\pi(N|\mathbf{T}, \mathbf{g}, \mathbf{m}, \boldsymbol{\phi}, \boldsymbol{\alpha}, \gamma_1, \gamma_2] = (\sum g_i + \gamma_1)/(1 + \gamma_2)$
end

4.2 A Dirichlet-Multinomial-Binomial Approach.

A second approach is both more computationally efficient and arguably more directly comparable to the sampling mechanism inherent in SAGE. For convenience, we begin by assuming that the total number of mRNA in the sample is $N \sim Pois(\lambda)$. Given this mRNA population size, the vector $\mathbf{g}$ of counts of each category of mRNA prior to tag formation follows a multinomial distribution,

$$[\mathbf{g}|\alpha] \sim \binom{N}{g_1, g_2, \ldots, g_k} m_1^{g_1} m_2^{g_2}, \ldots, m_k^{g_k}$$

whose probabilities are the m_i and are assumed to follow a Dirichlet distribution with parameter vector α.

The joint posterior distribution can now be written as,

$$\begin{aligned}[\mathbf{m}, \mathbf{g}, N|\mathbf{T}, \boldsymbol{\alpha}, \gamma_1, \gamma_2] \quad \propto \quad & \prod_{i=1}^{k} \binom{g_i}{t_i} \phi_i^{t_i} (1 - \phi_i)^{g_i - t_i} \frac{e^{-\lambda} \lambda^N}{N!} \\ & \times \binom{N}{g_1, g_2, \ldots, g_k} m_1^{g_1 + \alpha_1 - 1} m_2^{g_2 + \alpha_2 - 1}, \ldots, m_k^{g_k + \alpha_k - 1}.\end{aligned}$$