Gibbs sampling for B-LDA

贺成

Algorithm 1 Gibbs Sampling for B-LDA.

procedure GibbsSampling
    for each user $u = 1, \cdots, U$ **do**
        for u's n-th tweet, $n = 1, \cdots, N_u$ **do**
            Randomly assign a topic to $z_{u,n}$
            for each word $l = 1, \cdots, L_{u,n}$ **do**
                Randomly assign 0 or 1 to $y_{u,n,l}$
            end for
        end for
    end for
    for each Gibbs Sampling iteration **do**
        for each user $u = 1, \cdots, U$ **do**
            for u's n-th tweet, $n = 1, \cdots, N_u$ **do**
                Draw a topic $z_{u,n}$ according to Eqn. 2.1
                for each word $l = 1, \cdots, L_{u,n}$ **do**
                    Draw $y_{u,n,l}$ according to Eqn. 2.2
                end for
            end for
        end for
    end for
    Estimate model parameters θ, φ, ϕ', ϕ and ψ
end procedure

Notations	Descriptions
U	the total number of users
N_u	the total number of tweets in user u
$L_{u,n}$	the total number of words in u's n-th tweet
T	the total number of topics
b	a behavior in $\mathcal{B} = \{post, retweet, reply, mention\}$
y	a switch
z	a topic label
w	a word label
ϕ_t	topic-specific word distribution
ψ_t	topic-specific behavior distribution
ϕ'	background word distribution
θ_u	user-specific topic distribution
φ	Bernoulli distribution
α, η, β', β, γ	Dirichlet priors

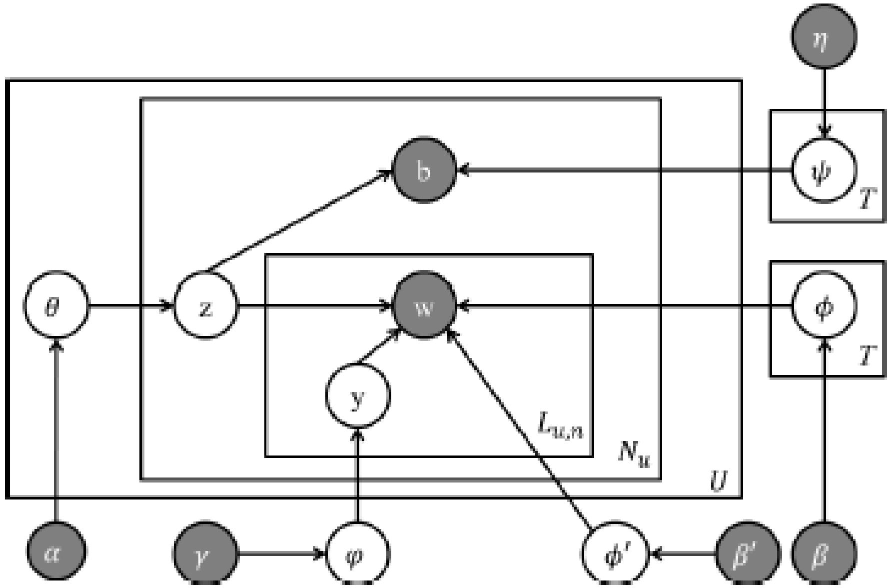


Figure 1: LDA-based behavior-topic model (B-LDA)

Hence, the problem is to compute the following two updating rules.

- To sample topic $z_{u,n}$:

$$p\left(z_{u,n} \middle| \vec{Z}_{\neg\{u,n\}}, \overrightarrow{W}, \vec{Y}, \vec{B}\right) = \frac{p\left(\vec{Z}, \overrightarrow{W}, \vec{Y}, \vec{B} \middle| \vec{\eta}, \vec{\beta}, \overrightarrow{\beta'}, \vec{\gamma}, \vec{\alpha}\right)}{p\left(\vec{Z}_{\neg\{u,n\}}, \overrightarrow{W}, \vec{Y}, \vec{B} | \vec{\eta}, \vec{\beta}, \overrightarrow{\beta'}, \vec{\gamma}, \vec{\alpha}\right)}$$

$$\propto \frac{p\left(\vec{Z}, \overrightarrow{W}, \vec{Y}, \vec{B} \middle| \vec{\eta}, \vec{\beta}, \overrightarrow{\beta'}, \vec{\gamma}, \vec{\alpha}\right)}{p\left(\vec{Z}_{\neg\{u,n\}}, \overrightarrow{W}, \vec{Y}, \vec{B}_{\neg\{u,n\}} | \vec{\eta}, \vec{\beta}, \overrightarrow{\beta'}, \vec{\gamma}, \vec{\alpha}\right)}$$

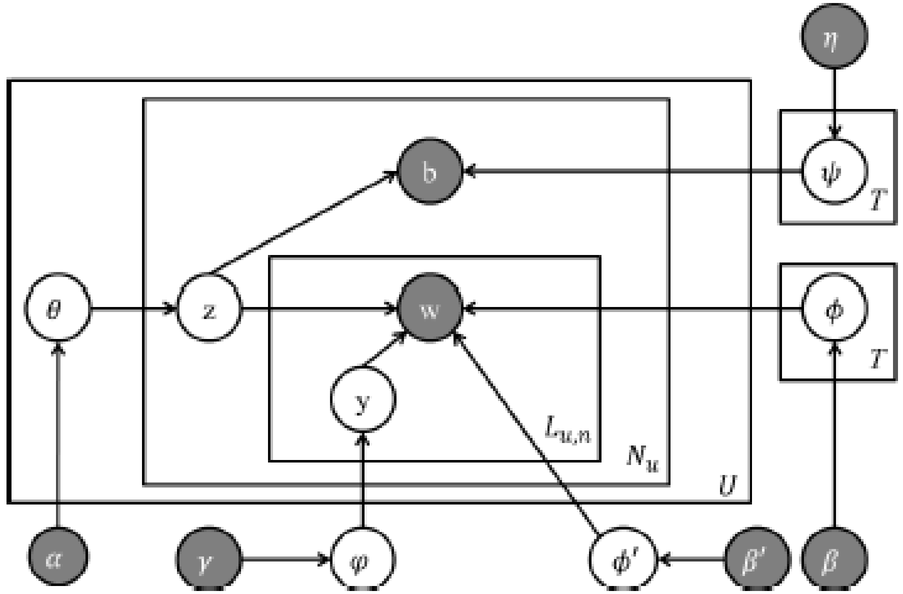


Figure 1: LDA-based behavior-topic model (B-LDA)

where $\vec{Z}_{\neg\{u,n\}}$ denotes the set of all the topics in the data sets not including the topic of user u's n-th tweet.

- To sample label $y_{u,n,l}$:

$$p\left(y_{u,n,l} \middle| \vec{Y}_{\neg\{u,n,l\}}, \vec{Z}, \overrightarrow{W}, \vec{B}\right) = \frac{p\left(\vec{Z}, \overrightarrow{W}, \vec{Y}, \vec{B} \middle| \vec{\eta}, \vec{\beta}, \overrightarrow{\beta'}, \vec{\gamma}, \vec{\alpha}\right)}{p\left(\vec{Y}_{\neg\{u,n,l\}}, \vec{Z}, \overrightarrow{W}, \vec{B} | \vec{\eta}, \vec{\beta}, \overrightarrow{\beta'}, \vec{\gamma}, \vec{\alpha}\right)}$$

$$\propto \frac{p\left(\vec{Z}, \overrightarrow{W}, \vec{Y}, \vec{B} \middle| \vec{\eta}, \vec{\beta}, \overrightarrow{\beta'}, \vec{\gamma}, \vec{\alpha}\right)}{p\left(\vec{Z}, \overrightarrow{W}_{\neg\{u,n,l\}}, \vec{Y}_{\neg\{u,n,l\}}, \vec{B} | \vec{\eta}, \vec{\beta}, \overrightarrow{\beta'}, \vec{\gamma}, \vec{\alpha}\right)}$$

□ Sampling topic $z_{u,n}$:

$$p(\vec{Z},\vec{W},\vec{Y},\vec{B}|\vec{\eta},\vec{\beta},\overrightarrow{\beta'},\vec{\gamma},\vec{\alpha}) = {\color{red}p(\vec{Z}|\vec{\alpha})}p(\vec{W}|\vec{Z},\vec{Y},\vec{\beta},\overrightarrow{\beta'})p(\vec{Y}|\vec{\gamma})p(\vec{B}|\vec{Z},\vec{\eta})$$

$$p(\vec{Z}|\vec{\alpha}) = \int p(\vec{Z}\,|\vec{\theta})p(\vec{\theta}|\vec{\alpha})d\vec{\theta} = \int \prod_{u=1}^{U}\prod_{t=1}^{T} p_t^{n_u^t} Dirichlet(\vec{\alpha})d\vec{\theta}_u$$

$$= \prod_{u=1}^{U}\int \prod_{t=1}^{T} p_t^{n_u^t} \frac{1}{\Delta(\vec{\alpha})}\prod_{t=1}^{T} p_t^{\alpha_t-1}\, d\vec{\theta}_u$$

$$= \prod_{u=1}^{U}\int \frac{1}{\Delta(\vec{\alpha})}\prod_{t=1}^{T} p_t^{n_u^t+\alpha_t-1} d\vec{\theta}_u$$

$$= \prod_{u=1}^{U}\frac{\Delta(\vec{n}_u +\ \vec{\alpha})}{\Delta(\vec{\alpha})}$$

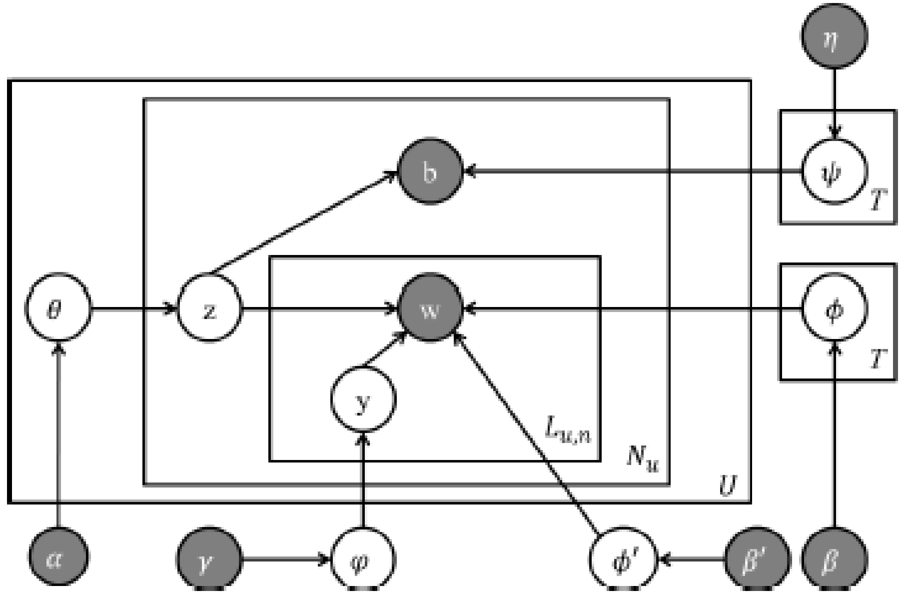


Figure 1: LDA-based behavior-topic model (B-LDA)

- Sampling topic $z_{u,n}$:

$$p(\vec{Z},\overrightarrow{W},\vec{Y},\vec{B}|\vec{\eta},\vec{\beta},\overrightarrow{\beta'},\vec{\gamma},\vec{\alpha}) = p(\vec{Z}|\vec{\alpha})p(\overrightarrow{W}|\vec{Z},\vec{Y},\vec{\beta},\overrightarrow{\beta'})p(\vec{Y}|\vec{\gamma})p(\vec{B}|\vec{Z},\vec{\eta})$$

$$p(\vec{B}|\vec{Z},\vec{\eta}) = \prod_{t=1}^{T}\int p(\vec{B}|\vec{Z}_t,\vec{\psi}_t)p(\vec{\psi}_t|\vec{\eta})d\vec{\psi}_t$$

$$= \prod_{t=1}^{T}\int \prod_{b=1}^{B} p_{t,b}^{n_t^b} Dirichlet(\vec{\eta})d\vec{\psi}_t$$

$$= \prod_{t=1}^{T}\int \frac{1}{\Delta(\vec{\eta})}\prod_{b=1}^{B} p_{t,b}^{n_t^b+\eta_b-1} d\vec{\psi}_t$$

$$= \prod_{t=1}^{T}\frac{\Delta(\vec{n}_t + \vec{\eta})}{\Delta(\vec{\eta})}$$

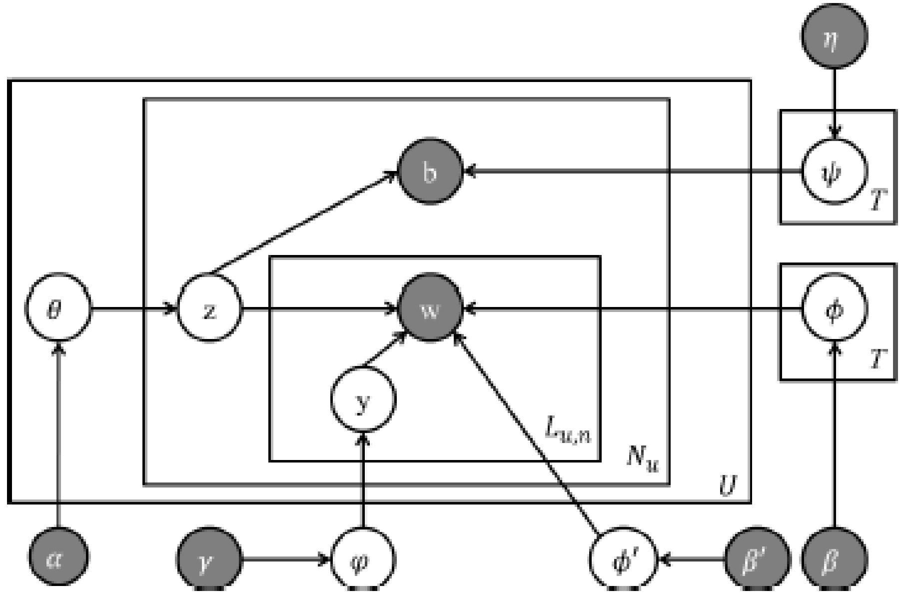


Figure 1: LDA-based behavior-topic model (B-LDA)

- Sampling topic $z_{u,n}$:

$$p(\vec{Z},\overrightarrow{W},\vec{Y},\vec{B}|\vec{\eta},\vec{\beta},\overrightarrow{\beta'},\vec{\gamma},\vec{\alpha}) = p(\vec{Z}|\vec{\alpha})p(\overrightarrow{W}|\vec{Z},\vec{Y},\vec{\beta},\overrightarrow{\beta'})p(\vec{Y}|\vec{\gamma})p(\vec{B}|\vec{Z},\vec{\eta})$$

$$p(\vec{Y}|\vec{\gamma}) = \int p(\vec{Y}\,|\vec{\varphi})p(\vec{\varphi}|\vec{\gamma})d\vec{\varphi} = \int \prod_{y=0}^{1} p_y^{n^y} Dirichlet(\vec{\gamma})d\vec{\varphi}$$

$$= \int \prod_{y=0}^{1} p_y^{n^y} \frac{1}{\Delta(\vec{\gamma})} \prod_{y=0}^{1} p_y^{\gamma_y - 1}\, d\vec{\varphi}$$

$$= \int \frac{1}{\Delta(\vec{\gamma})} \prod_{y=0}^{1} p_y^{n^y+\gamma_y-1} d\vec{\varphi}$$

$$= \frac{\Delta(n_{(\cdot)}^{y} + \vec{\gamma})}{\Delta(\vec{\gamma})}$$

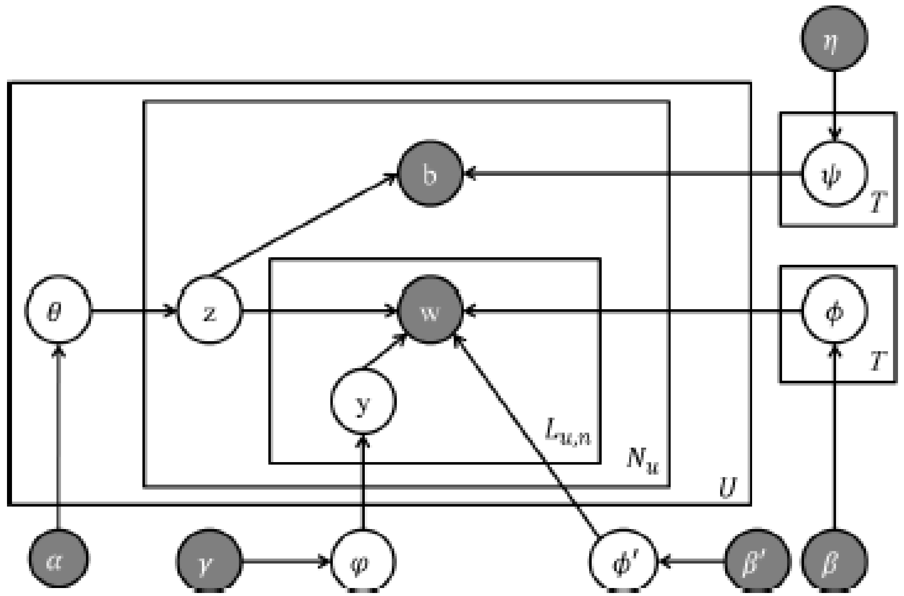


Figure 1: LDA-based behavior-topic model (B-LDA)

◻ Sampling topic $z_{u,n}$:

$$p(\vec{Z},\vec{W},\vec{Y},\vec{B}|\vec{\eta},\vec{\beta},\vec{\beta'},\vec{\gamma},\vec{\alpha}) = p(\vec{Z}|\vec{\alpha})p(\vec{W}|\vec{Z},\vec{Y},\vec{\beta},\vec{\beta'})p(\vec{Y}|\vec{\gamma})p(\vec{B}|\vec{Z},\vec{\eta})$$

We assume each word has a corresponding label y that indicates which model it is sampled from. Specifically, if $y = 0$, the word is sampled from the background model; if $y = 1$, it is from a topic specific model. To derive $p(\vec{W}|\vec{Z},\vec{Y},\vec{\beta},\vec{\beta'})$,we then need to consider two types of word distributions ϕ and ϕ'.

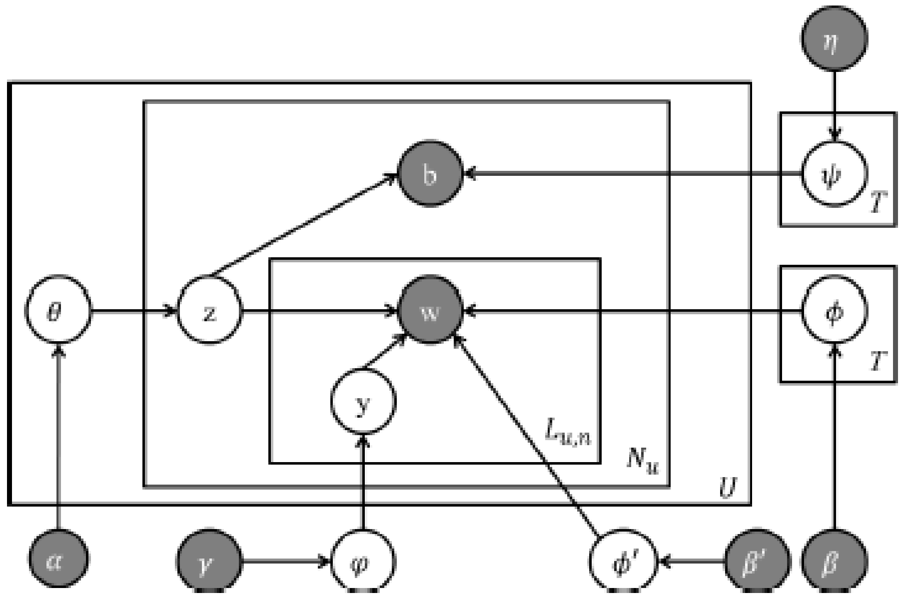


Figure 1: LDA-based behavior-topic model (B-LDA)

$$p(\vec{W}|\vec{Z},\vec{Y},\vec{\beta},\vec{\beta'}) = \int\int p(\vec{W}|\vec{Z},\vec{Y},\vec{\phi},\vec{\phi}')p(\vec{\phi}|\vec{\beta})p(\vec{\phi}'|\vec{\beta}')d\vec{\phi}'d\vec{\phi}$$

$$= \int\int \prod_{w=1}^{V} p_{w,y=0}^{n_{y=0}^{w}} \prod_{t=1}^{T}\prod_{w=1}^{V} p_{t,w,y=1}^{n_{t,y=1}^{w}} \frac{1}{\Delta(\beta)} Dirichlet(\vec{\beta}) \frac{1}{\Delta(\vec{\beta}')} Dirichlet(\vec{\beta}')d\vec{\phi}'d\vec{\phi}$$

$$= \int\int \prod_{w=1}^{V} p_{w,y=0}^{n_{y=0}^{w}} \prod_{t=1}^{T}\prod_{w=1}^{V} p_{t,w,y=1}^{n_{t,y=1}^{w}} \frac{1}{\Delta(\vec{\beta})} \prod_{w=1}^{V} p_{t,w,y=1}^{\beta^{w}-1} \frac{1}{\Delta(\vec{\beta}')} \prod_{w=1}^{V} p_{w,y=0}^{\beta'^{w}-1}\ d\vec{\phi}'d\vec{\phi}$$

$$= \int\int \frac{1}{\Delta(\vec{\beta})} \prod_{t=1}^{T}\prod_{w=1}^{V} p_{t,w,y=1}^{n_{t,y=1}^{w}+\beta^{w}-1} \frac{1}{\Delta(\vec{\beta}')} \prod_{w=1}^{V} p_{w,y=0}^{n_{y=0}^{w}+\beta'^{w}-1}\ d\vec{\phi}'d\vec{\phi}$$

$$= \frac{\Delta(n_{y=0}^{w} + \beta')}{\Delta(\vec{\beta}')} \prod_{t=1}^{T} \frac{\Delta(n_{t,y=1}^{w} + \beta)}{\Delta(\vec{\beta})}$$

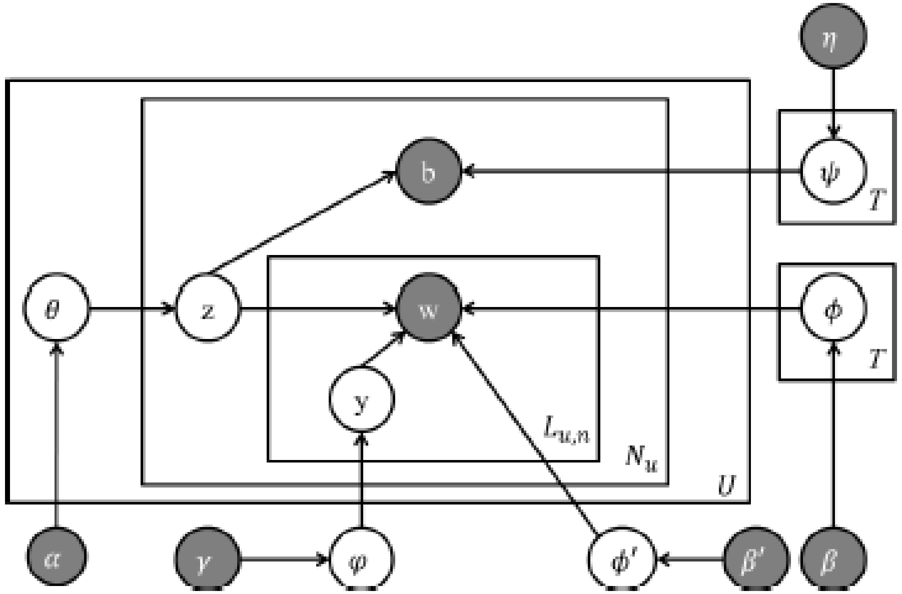


Figure 1: LDA-based behavior-topic model (B-LDA)

因此：

$$p(\vec{Z},\overrightarrow{W},\vec{Y},\vec{B}|\vec{\eta},\vec{\beta},\overrightarrow{\beta'},\vec{\gamma},\vec{\alpha})$$

$$= p(\vec{Y}|\vec{\gamma})p(\overrightarrow{W}|\vec{Z},\vec{Y},\vec{\beta},\overrightarrow{\beta'})p(\vec{B}|\vec{Z},\vec{\eta})p(\vec{Z}|\vec{\alpha})$$

$$= \frac{\Delta(n^{y}_{(\cdot)} + \vec{\gamma})}{\Delta(\vec{\gamma})} \cdot \frac{\Delta(n^{w}_{y=0} + \beta')}{\Delta(\vec{\beta}')} \prod_{t=1}^{T} \frac{\Delta(n^{w}_{t,y=1} + \beta)}{\Delta(\vec{\beta})} \cdot \prod_{t=1}^{T} \frac{\Delta(\vec{n}^{b}_{t} + \vec{\eta})}{\Delta(\vec{\eta})} \cdot \prod_{u=1}^{U} \frac{\Delta(\vec{n}^{t}_{u} + \vec{\alpha})}{\Delta(\vec{\alpha})}$$

$let\ c\ denote\ \{u, n\}$:

$$p\left(\vec{Z}_c \middle| \vec{Z}_{\neg c}, \overrightarrow{W}, \vec{Y}, \vec{B}\right) = \frac{p\left(\vec{Z}, \overrightarrow{W}, \vec{Y}, \vec{B} \middle| \vec{\eta}, \vec{\beta}, \overrightarrow{\beta'}, \vec{\gamma}, \vec{\alpha}\right)}{p\left(\vec{Z}_{\neg c}, \overrightarrow{W}, \vec{Y}, \vec{B} | \vec{\eta}, \vec{\beta}, \overrightarrow{\beta'}, \vec{\gamma}, \vec{\alpha}\right)}$$

$$\propto \frac{p\left(\vec{Z}, \overrightarrow{W}, \vec{Y}, \vec{B} \middle| \vec{\eta}, \vec{\beta}, \overrightarrow{\beta'}, \vec{\gamma}, \vec{\alpha}\right)}{p\left(\vec{Z}_{\neg c}, \overrightarrow{W}, \vec{Y}, \vec{B}_{\neg c} | \vec{\eta}, \vec{\beta}, \overrightarrow{\beta'}, \vec{\gamma}, \vec{\alpha}\right)}$$

$$= \frac{\cancel{\frac{\Delta(n^{\boldsymbol{y}}_{(\cdot)} + \vec{\gamma})}{\Delta(\vec{\gamma})}} \cdot \cancel{\frac{\Delta\left(n^{w}_{y=0} + \beta'\right)}{\Delta\left(\vec{\beta}'\right)}} \prod_{t=1}^{T} \frac{\Delta\left(n^{w}_{t,y=1} + \beta\right)}{\cancel{\Delta(\vec{\beta})}} \cdot \prod_{t=1}^{T} \frac{\Delta\left(\vec{n}^{b}_{t} + \vec{\eta}\right)}{\cancel{\Delta(\vec{\eta})}} \cdot \prod_{u=1}^{U} \frac{\Delta(\vec{n}^{t}_{u} + \vec{\alpha})}{\cancel{\Delta(\vec{\alpha})}}}{\cancel{\frac{\Delta(n^{\boldsymbol{y}}_{(\cdot)} + \vec{\gamma})}{\Delta(\vec{\gamma})}} \cdot \cancel{\frac{\Delta\left(n^{w}_{y=0} + \beta'\right)}{\Delta\left(\vec{\beta}'\right)}} \prod_{t=1}^{T} \frac{\Delta\left(n^{w}_{t,y=1,\neg c} + \beta\right)}{\cancel{\Delta(\vec{\beta})}} \cdot \prod_{t=1}^{T} \frac{\Delta\left(\vec{n}^{b}_{t,\neg c} + \vec{\eta}\right)}{\cancel{\Delta(\vec{\eta})}} \cdot \prod_{u=1}^{U} \frac{\Delta\left(\vec{n}^{t}_{u,\neg c} + \vec{\alpha}\right)}{\cancel{\Delta(\vec{\alpha})}}}$$

$$= \prod_{t=1}^{T} \frac{\Delta\left(n^{w}_{t,y=1} + \vec{\beta}\right)}{\Delta\left(n^{w}_{t,y=1,\neg c} + \vec{\beta}\right)} \cdot \prod_{t=1}^{T} \frac{\Delta\left(\vec{n}^{b}_{t} + \vec{\eta}\right)}{\Delta\left(\vec{n}^{b}_{t,\neg c} + \vec{\eta}\right)} \cdot \prod_{u=1}^{U} \frac{\Delta(\vec{n}^{t}_{u} + \vec{\alpha})}{\Delta\left(\vec{n}^{t}_{u,\neg c} + \vec{\alpha}\right)}$$

$$\Rightarrow p\left(\vec{z}_c = z \middle| \vec{Z}_{\neg c}, \overrightarrow{W}, \vec{Y}, \vec{B}\right) = \frac{\Delta\left(n^{w}_{t,y=1} + \vec{\beta}\right)}{\Delta\left(n^{w}_{t,y=1,\neg c} + \vec{\beta}\right)} \cdot \frac{\Delta\left(\vec{n}^{b}_{z} + \vec{\eta}\right)}{\Delta\left(\vec{n}^{b}_{z,\neg c} + \vec{\eta}\right)} \cdot \frac{\Delta\left(\vec{n}^{t}_{u} + \vec{\alpha}\right)}{\Delta\left(\vec{n}^{t}_{u,\neg c} + \vec{\alpha}\right)}$$

$$\frac{\Delta\left(n_{t,y=1}^{w}+\vec{\beta}\right)}{\Delta\left(n_{t,y=1,\neg c}^{w}+\vec{\beta}\right)}=\frac{\frac{\prod_{w=1}^{V}\Gamma\left(n_{z,y=1}^{w}+\beta\right)}{\Gamma\left(\sum_{w=1}^{V}n_{z,y=1}^{w}+V\beta\right)}}{\frac{\prod_{w=1}^{V}\Gamma\left(n_{z,y=1,\neg c}^{w}+\beta\right)}{\Gamma\left(\sum_{w=1}^{V}n_{z,y=1,\neg c}^{w}+V\beta\right)}}$$

$$=\frac{\prod_{w=1}^{V}\Gamma\left(n_{z,y=1}^{w}+\beta\right)}{\prod_{w=1}^{V}\Gamma\left(n_{z,y=1,\neg c}^{w}+\beta\right)}\cdot\frac{\Gamma\left(\sum_{w=1}^{V}n_{z,y=1,\neg c}^{w}+V\beta\right)}{\Gamma\left(\sum_{w=1}^{V}n_{z,y=1}^{w}+V\beta\right)} \tag{2}$$

$$=\frac{\prod_{w=1}^{V}\prod_{i=1}^{n_{c,y=1}^{w}}\left(n_{z,y=1,\neg c}^{w}+\beta+i-1\right)}{\prod_{j=1}^{n_{c,y=1}^{W}}\left(\sum_{w=1}^{V}n_{z,y=1,\neg c}^{w}+V\beta+j-1\right)} \tag{3}$$

注: (2)->(3):

$$\frac{\prod_{w=1}^{V}\Gamma\left(n_{z,y=1}^{w}+\beta\right)}{\prod_{w=1}^{V}\Gamma\left(n_{z,y=1,\neg c}^{w}+\beta\right)}=\frac{\Gamma\left(n_{z,y=1}^{1}+\beta\right)\cdot\Gamma\left(n_{z,y=1}^{2}+\beta\right)\cdots\Gamma\left(n_{z,y=1}^{V}+\beta\right)}{\Gamma\left(n_{z,y=1,\neg c}^{1}+\beta\right)\cdot\Gamma\left(n_{z,y=1,\neg c}^{2}+\beta\right)\cdots\Gamma\left(n_{z,y=1,\neg c}^{V}+\beta\right)}$$

$$=\prod_{w=1}^{V}\prod_{i=1}^{n_{c,y=1}^{w}}\left(n_{z,y=1,\neg c}^{w}+\beta+i-1\right)$$

where $n_{c,y=1}^{w}$ denotes the number of times word w occurs as topical words ;

$n_{c,y=1}^{W}$ is the total number of topical words in user $u's$ $n-th$ tweets.

$$\frac{\Delta\left(\vec{n}_z^b + \vec{\eta}\right)}{\Delta\left(\vec{n}_{z,\neg c}^b + \vec{\eta}\right)} = \frac{\frac{\prod_{b=1}^{B}\Gamma(n_z^b + \eta)}{\Gamma\left(\sum_{b=1}^{B} n_z^b + B\eta\right)}}{\frac{\prod_{b=1}^{B}\Gamma(n_{z,\neg c}^b + \eta)}{\Gamma\left(\sum_{b=1}^{B} n_{z,\neg c}^b + B\eta\right)}} = \frac{\prod_{b=1}^{B}\Gamma(n_z^b + \eta)}{\prod_{b=1}^{B}\Gamma(n_{z,\neg c}^b + \eta)} \frac{\Gamma\left(\sum_{b=1}^{B} n_{z,\neg c}^b + B\eta\right)}{\Gamma\left(\sum_{b=1}^{B} n_z^b + B\eta\right)}$$

$$= \frac{\Gamma(n_z^b + \eta)}{\Gamma(n_{z,\neg c}^b + \eta)} \frac{1}{\sum_{b=1}^{B} n_{z,\neg c}^b + B\eta}$$

$$= \frac{n_{z,\neg c}^b + \eta}{\sum_{b=1}^{B} n_{z,\neg c}^b + B\eta}$$

同理：$\frac{\Delta\left(\vec{n}_u^t + \vec{\alpha}\right)}{\Delta\left(\vec{n}_{u,\neg c}^t + \vec{\alpha}\right)} = \frac{n_{u,\neg c}^z + \alpha}{\sum_{\mathrm{t}=1}^{T} n_{u,\neg c}^z + T\alpha}$

where $n_{z,\neg c}^b$ denotes number of times topic z co-occurs with behavior b without considering the current tweet, $n_{u,\neg c}^z$ denotes number of times topic z is sampled in user $u's$ tweets without considering the current tweet.

因此，

$$p\left(\vec{z}_c = z \middle| \vec{Z}_{\neg c}, \vec{W}, \vec{Y}, \vec{B}\right) = \frac{\prod_{w=1}^{V} \prod_{i=1}^{n_{c,y=1}^{w}} (n_{z,y=1,\neg c}^{w} + \beta + i - 1)}{\prod_{j=1}^{n_{c,y=1}^{w}} (\sum_{w=1}^{V} n_{z,y=1,\neg c}^{w} + V\beta + j - 1)} \cdot \frac{n_{z,\neg c}^{b} + \eta}{\sum_{b=1}^{B} n_{z,\neg c}^{b} + B\eta} \cdot \frac{n_{u,\neg c}^{z} + \alpha}{\sum_{\mathrm{t}=1}^{T} n_{u,\neg c}^{z} + T\alpha}$$

■ To sample label $y_{u,n,l}$:

Let d be $\{u, n, l\}$,

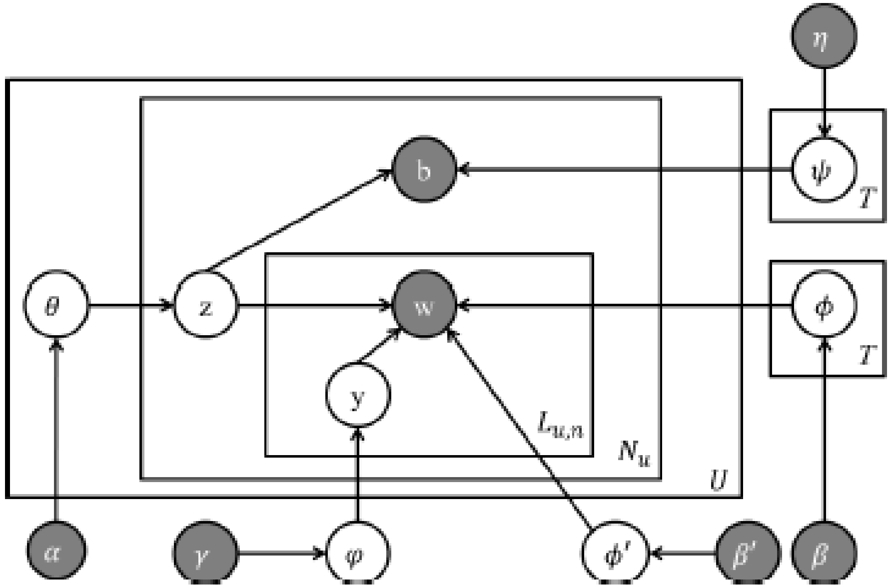


Figure 1: LDA-based behavior-topic model (B-LDA)

$$p\left(y_d \middle| \vec{Y}_{\neg d}, \vec{Z}, \vec{W}, \vec{B}\right) = \frac{p\left(\vec{Z}, \vec{W}, \vec{Y}, \vec{B} \middle| \vec{\eta}, \vec{\beta}, \overrightarrow{\beta'}, \vec{\gamma}, \vec{\alpha}\right)}{p\left(\vec{Y}_{\neg d}, \vec{Z}, \vec{W}, \vec{B} \middle| \vec{\eta}, \vec{\beta}, \overrightarrow{\beta'}, \vec{\gamma}, \vec{\alpha}\right)}$$

$$\propto \frac{p\left(\vec{Z}, \vec{W}, \vec{Y}, \vec{B} \middle| \vec{\eta}, \vec{\beta}, \overrightarrow{\beta'}, \vec{\gamma}, \vec{\alpha}\right)}{p\left(\vec{Z}, \vec{W}_{\neg d}, \vec{Y}_{\neg d}, \vec{B} \middle| \vec{\eta}, \vec{\beta}, \overrightarrow{\beta'}, \vec{\gamma}, \vec{\alpha}\right)}$$

$$= \frac{\frac{\Delta(n^{y}_{(\cdot)} + \vec{\gamma})}{\cancel{\Delta(\vec{\gamma})}} \cdot \frac{\Delta\left(n^{w}_{y=0} + \vec{\beta}'\right)}{\cancel{\Delta(\vec{\beta}')}} \prod_{t=1}^{T} \frac{\Delta\left(n^{w}_{t,y=1} + \vec{\beta}\right)}{\cancel{\Delta(\vec{\beta})}} \cdot \prod_{t=1}^{T} \cancel{\frac{\Delta\left(\vec{n}^{b}_{t} + \vec{\eta}\right)}{\Delta(\vec{\eta})}} \cdot \prod_{u=1}^{U} \cancel{\frac{\Delta\left(\vec{n}^{t}_{u} + \vec{\alpha}\right)}{\Delta(\vec{\alpha})}}}{\frac{\Delta(n^{y}_{\neg d} + \vec{\gamma})}{\cancel{\Delta(\vec{\gamma})}} \cdot \frac{\Delta\left(n^{w}_{y=0,\neg d} + \vec{\beta}'\right)}{\cancel{\Delta(\vec{\beta}')}} \prod_{t=1}^{T} \frac{\Delta\left(n^{w}_{t,y=1,\neg d} + \vec{\beta}\right)}{\cancel{\Delta(\vec{\beta})}} \cdot \prod_{t=1}^{T} \cancel{\frac{\Delta\left(\vec{n}^{b}_{t} + \vec{\eta}\right)}{\Delta(\vec{\eta})}} \cdot \prod_{u=1}^{U} \cancel{\frac{\Delta\left(\vec{n}^{t}_{u} + \vec{\alpha}\right)}{\Delta(\vec{\alpha})}}}$$

$$= \frac{\Delta(n^{y}_{(\cdot)} + \vec{\gamma})}{\Delta(n^{y}_{\neg d} + \vec{\gamma})} \cdot \frac{\Delta\left(n^{w}_{y=0} + \vec{\beta}'\right)}{\Delta\left(n^{w}_{y=0,\neg d} + \vec{\beta}'\right)} \prod_{t=1}^{T} \frac{\Delta\left(n^{w}_{t,y=1} + \vec{\beta}\right)}{\Delta\left(n^{w}_{t,y=1,\neg d} + \vec{\beta}\right)}$$

$$p\left(y_d = 0 \middle| \vec{Y}_{\neg d}, \vec{Z}, \vec{W}, \vec{B}\right) = \frac{\Delta(n^{y}_{(.)} + \vec{\gamma})}{\Delta(n^{y}_{\neg d} + \vec{\gamma})} \cdot \frac{\Delta\left(n^{w}_{y=0} + \vec{\beta}'\right)}{\Delta\left(n^{w}_{y=0,\neg d} + \vec{\beta}'\right)}$$

$$\frac{\Delta(n^{y}_{(.)} + \vec{\gamma})}{\Delta(n^{y}_{\neg d} + \vec{\gamma})} = \frac{\dfrac{\prod_{y=0}^{1} \Gamma(n^{y}_{(.)} + \gamma)}{\Gamma(\sum_{y=0}^{1} n^{y}_{(.)} + 2\gamma)}}{\dfrac{\prod_{y=0}^{1} \Gamma(n^{y}_{\neg d} + \gamma)}{\Gamma(\sum_{y=0}^{1} n^{y}_{\neg d} + 2\gamma)}}$$

$$= \frac{\prod_{y=0}^{1} \Gamma(n^{y}_{(.)} + \gamma)}{\prod_{y=0}^{1} \Gamma(n^{y}_{\neg d} + \gamma)} \cdot \frac{1}{\sum_{y=0}^{1} n^{y}_{\neg d} + 2\gamma}$$

$$= \frac{n^{y_d=0}_{\neg d} + \gamma}{\sum_{y=0}^{1} n^{y}_{\neg d} + 2\gamma}$$

$$\frac{\Delta\left(n^{w}_{y=0} + \vec{\beta}'\right)}{\Delta\left(n^{w}_{y=0,\neg d} + \vec{\beta}'\right)} = \frac{\dfrac{\prod_{w=1}^{V} \Gamma(n^{w}_{y=0} + \beta')}{\Gamma(\sum_{w=1}^{V} n^{w}_{y=0} + V\beta')}}{\dfrac{\prod_{w=1}^{V} \Gamma(n^{w}_{y=0,\neg d} + \beta')}{\Gamma(\sum_{w=1}^{V} n^{w}_{y=0,\neg d} + V\beta')}}$$

$$= \frac{n^{w_d}_{y=0,\neg d} + \beta'}{\sum_{w=1}^{V} n^{w}_{y=0,\neg d} + V\beta'}$$

$$\Rightarrow p\left(y_d = 0 \middle| \vec{Y}_{\neg d}, \vec{Z}, \vec{W}, \vec{B}\right) = \frac{n^{y_d=0}_{\neg d} + \gamma}{\sum_{y=0}^{1} n^{y}_{\neg d} + 2\gamma} \cdot \frac{n^{w_d}_{y=0,\neg d} + \beta'}{\sum_{w=1}^{V} n^{w}_{y=0,\neg d} + V\beta'}$$

同理：$p\left(y_d=1 \middle| \vec{Y}_{\neg d}, \vec{Z}, \vec{W}, \vec{B}\right) = \frac{n_{\neg d}^{y_d=1} + \gamma}{\sum_{y=0}^{1} n_{\neg d}^{y} + 2\gamma} \cdot \frac{n_{z_c, y=1, \neg d}^{w_d} + \beta}{\sum_{w=1}^{V} n_{z_c, y=1, \neg d}^{w} + V\beta}$

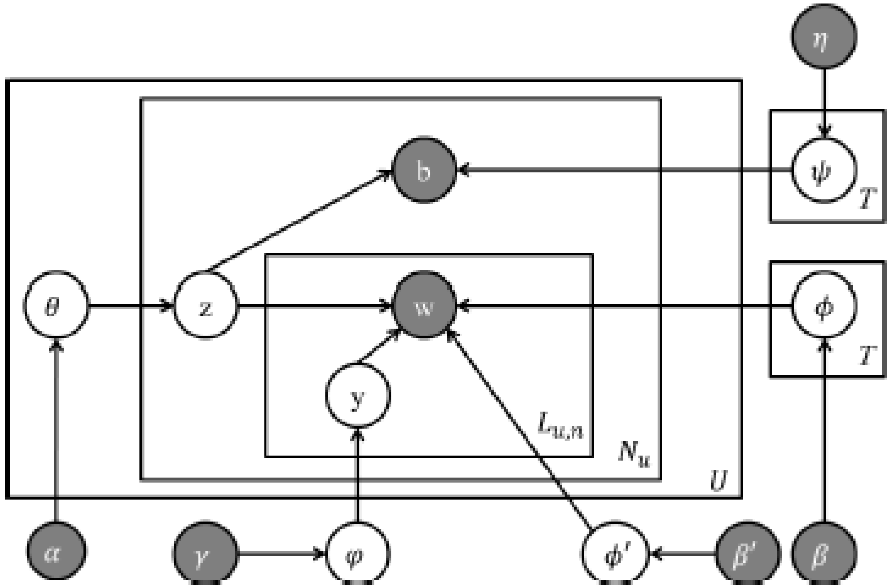


Figure 1: LDA-based behavior-topic model (B-LDA)

参数估计：

$$\phi'_w = \frac{n_{y=0}^{w} + \beta'}{\sum_{w=1}^{V} n_{y=0}^{w} + V\beta'}$$

$$\phi_{t,w} = \frac{n_{t,y=1}^{w} + \beta}{\sum_{w=1}^{V} n_{t,y=1}^{w} + V\beta}$$

$$\psi_{t,b} = \frac{n_t^b + \eta}{\sum_{b=1}^{B} n_t^b + B\eta}$$

$$\varphi_y = \frac{n_{(.)}^{y} + \gamma}{\sum_{y=0}^{1} n_{(.)}^{y} + 2\gamma}$$

$$\theta_{u,t} = \frac{n_u^t + \alpha}{\sum_{t=1}^{T} n_u^t + T\alpha}$$

where $n_{y=0}^{w}$ is the number of times w appears as background word, $n_{t,y=1}^{w}$ is the number of times w is sampled as topical word specific to topic t, n_t^b is number of time posting behavior b co-occurs with topic t, $n_{(.)}^{y}$ is number of times y appears, where n_u^t is, when given the user u, number of times t is sampled.

关于TOIS这篇：
关于这部分推导：

$$
\begin{aligned}
&p(z_{u,s} = a | \mathbf{z}_{\neg\{u,s\}}, \mathbf{y}, \mathbf{l}, \boldsymbol{\Delta}) \\
&= \frac{p(z_{u,s} = a, \mathbf{z}_{\neg\{u,s\}}, \mathbf{y}, \mathbf{l}, \boldsymbol{\Delta})}{p(\mathbf{z}_{\neg\{u,s\}}, \mathbf{y}, \mathbf{l}, \boldsymbol{\Delta})} \\
&= \frac{p(z_{u,s} = a, \mathbf{z}_{\neg\{u,s\}}, \mathbf{y}, \mathbf{l}, \boldsymbol{\Delta})}{p(\mathbf{z}_{\neg\{u,s\}}, \mathbf{y}_{\neg\{u,s\}}, \mathbf{l}_{\neg\{u,s\}}, \boldsymbol{\Delta}_{\neg\{u,s\}})} \cdot \frac{1}{p(\mathbf{y}_{\{u,s\}}, \boldsymbol{\Delta}_{\{u,s\}}, \mathbf{l}_{\{u,s\}} | \mathbf{z}_{\neg\{u,s\}}, \mathbf{y}_{\neg\{u,s\}}, \mathbf{l}_{\neg\{u,s\}}, \boldsymbol{\Delta}_{\neg\{u,s\}})} \\
&\propto \frac{p(z_{u,s} = a, \mathbf{z}_{\neg\{u,s\}}, \mathbf{y}, \mathbf{l}, \boldsymbol{\Delta})}{p(\mathbf{z}_{\neg\{u,s\}}, \mathbf{y}_{\neg\{u,s\}}, \mathbf{l}_{\neg\{u,s\}}, \boldsymbol{\Delta}_{\neg\{u,s\}})} \qquad (1)
\end{aligned}
$$

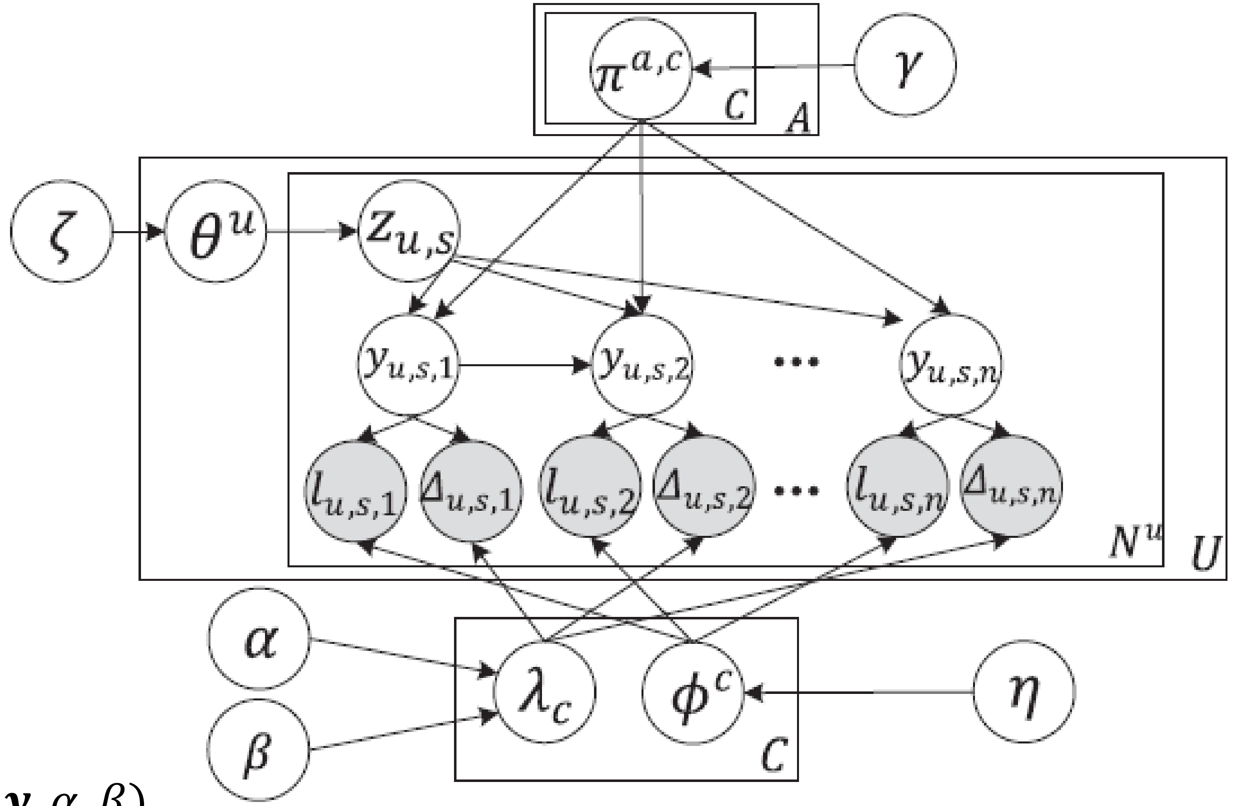
$$
\begin{aligned}
&= \frac{p(z_{u,s} = a, \mathbf{z}_{\neg\{u,s\}})}{p(\mathbf{z}_{\neg\{u,s\}})} \\
&\quad \times \frac{p(\mathbf{y} | z_{u,s} = a, \mathbf{z}_{\neg\{u,s\}})}{p(\mathbf{y}_{\neg\{u,s\}} | \mathbf{z}_{\neg\{u,s\}})} \\
&\quad \times \frac{p(\mathbf{l} | z_{u,s} = a, \mathbf{z}_{\neg\{u,s\}}, \mathbf{y})}{p(\mathbf{l}_{\neg\{u,s\}} | \mathbf{z}_{\neg\{u,s\}}, \mathbf{y}_{\neg\{u,s\}})} \\
&\quad \times \frac{p(\boldsymbol{\Delta} | z_{u,s} = a, \mathbf{z}_{\neg\{u,s\}}, \mathbf{y})}{p(\boldsymbol{\Delta}_{\neg\{u,s\}} | \mathbf{z}_{\neg\{u,s\}}, \mathbf{y}_{\neg\{u,s\}})}, \qquad (2)
\end{aligned}
$$

我自己的想法是：

$$p\left(z_{u,s}=a\middle|\boldsymbol{z}_{\neg\{u,s\}},\boldsymbol{y},\boldsymbol{l},\Delta\right)$$

$$=\frac{p(\boldsymbol{z},\boldsymbol{y},\boldsymbol{l},\Delta|\zeta,\alpha,\beta,\eta,\gamma)}{p\left(\boldsymbol{z}_{\neg\{u,s\}},\boldsymbol{y},\boldsymbol{l},\Delta\middle|\zeta,\alpha,\beta,\eta,\gamma\right)}$$

$$\propto\frac{p(\boldsymbol{z},\boldsymbol{y},\boldsymbol{l},\Delta|\zeta,\alpha,\beta,\eta,\gamma)}{p\left(\boldsymbol{z}_{\neg\{u,s\}},\boldsymbol{y}_{\neg\{u,s\}},\boldsymbol{l}_{\neg\{u,s\}},\boldsymbol{\Delta}_{\neg\{u,s\}}\middle|\zeta,\alpha,\beta,\eta,\gamma\right)}$$



$$p(\boldsymbol{z},\boldsymbol{y},\boldsymbol{l},\Delta|\zeta,\alpha,\beta,\eta,\gamma)=p(\boldsymbol{z}|\zeta)p(\boldsymbol{y}|\boldsymbol{z},\gamma)p(\boldsymbol{l}|\boldsymbol{y},\eta)p(\boldsymbol{\Delta}|\boldsymbol{y},\alpha,\beta)$$

注：关于这个式子，我自己根据盘子图是这样想的，也不知道对不对

$$P(Z|\xi) = \int P(Z|\theta)\, P(\theta|\xi)\, d\theta$$

$$= \int \prod_{u=1}^{U} \prod_{a=1}^{A} P_{u,a}^{N_a^u} \cdot \text{Dirichlet}(\xi)\, d\theta$$

$$= \int \prod_{u=1}^{U} \prod_{a=1}^{A} P_{u,a}^{N_a^u + \xi_a - 1} \cdot \frac{1}{\Delta(\xi)}\, d\theta$$

$$= \prod_{u=1}^{U} \frac{\Delta(N_a^u + \xi)}{\Delta(\xi)}$$

$$P(Z_{u,s} = a \mid Z_{\neg(u,s)}) \overset{\cdot \cdot}{=} \frac{P(Z)}{P(Z_{\neg(u,s)})} = \frac{P(Z|\xi)}{P(Z_{\neg(u,s)}|\xi)} = \frac{\prod_{u=1}^{U} \frac{\Delta(N_a^u + \xi)}{\Delta(\xi)}}{\prod_{u=1}^{U} \frac{\Delta(N_{a,\neg(u,s)}^u + \xi)}{\Delta(\xi)}} = \frac{\Delta(N_a^u + \xi)}{\Delta(N_{a,\neg(u,s)}^u + \xi)}$$

$$\frac{\Delta(N_a^u + \xi)}{\Delta(N_{a,\neg(u,s)}^u + \xi)} = \frac{\dfrac{\prod_{a=1}^{A} \Gamma(N_a^u + \xi)}{\Gamma(\sum_{a=1}^{A} N_a^u + A\xi)}}{\dfrac{\prod_{a=1}^{A} \Gamma(N_{a,\neg(u,s)}^u + \xi)}{\Gamma(\sum_{a=1}^{A} N_{a,\neg(u,s)}^u + A\xi)}} = \frac{N_{a,\neg(u,s)}^u + \xi}{\sum_{a=1}^{A} N_{a,\neg(u,s)}^u + A\xi} \propto \frac{N_a^u + \xi}{N_{(\cdot)}^u + A\xi}$$

$$\frac{p(\mathbf{y}|z_{u,s}=a,\mathbf{z}_{\neg\{u,s\}})}{p(\mathbf{y}_{\neg\{u,s\}}|\mathbf{z}_{\neg\{u,s\}})} \propto \prod_{c_1=0}^{C}\left\{\frac{\Gamma(N_{(\cdot)}^{a,c_1}+C\gamma)}{\Gamma(N_{(\cdot)}^{a,c_1}+N_{(\cdot)}^{u,s,c_1}+C\gamma)}\prod_{c_2=1}^{C}\frac{\Gamma(N_{c_2}^{a,c_1}+N_{c_2}^{c_1}+\gamma)}{\Gamma(N_{c_2}^{t,c_1}+\gamma)}\right\}, \qquad (2)$$

$$\frac{p(\mathbf{l}|z_{u,s}=a,\mathbf{z}_{\neg\{u,s\}},\mathbf{y})}{p(\mathbf{l}_{\neg\{u,s\}}|\mathbf{z}_{\neg\{u,s\}},\mathbf{y}_{\neg\{u,s\}})} \propto \prod_{c=1}^{C}\left\{\frac{\Gamma(N_{(\cdot)}^{c}+L\eta)}{\Gamma(N_{(\cdot)}^{c}+N_{(\cdot)}^{u,s,c}+L\eta)}\prod_{l=1}^{L}\frac{\Gamma(N_{l}^{c}+N_{l}^{u,s,c}+\eta)}{\Gamma(N_{l}^{c}+\eta)}\right\}, \qquad (3)$$

$$\frac{p(\boldsymbol{\Delta}|z_{u,s}=a,\mathbf{z}_{\neg\{u,s\}},\mathbf{y})}{p(\boldsymbol{\Delta}_{\neg\{u,s\}}|\mathbf{z}_{\neg\{u,s\}},\mathbf{y}_{\neg\{u,s\}})} \propto \prod_{c=1}^{C}\lambda_c^{N_{(\cdot)}^{u,s,c}}e^{-\lambda_c N^{u,s}\overline{\Delta^{u,s}}}. \qquad (4)$$

但是这三个式子我尝试了很多遍还是没推导出来，主要感觉奇怪的地方是：比如（2）式，按照基本LDA，Twitter-LDA，B-LDA的思路，我觉得不应该再出现gamma函数了，因为就像第一个式子那种应该可以约掉的。还有就是$(z_{u,s}=a,z_{\neg\{u,s\}})$应该就是Z吧？根据盘子图，l在给定y的时候我认为就和z无关了，因此第（3）个式子中，$p(l|z,y)$中z应该可以忽略，就为$p(l|y)$。第（4）个也有相同的疑惑。