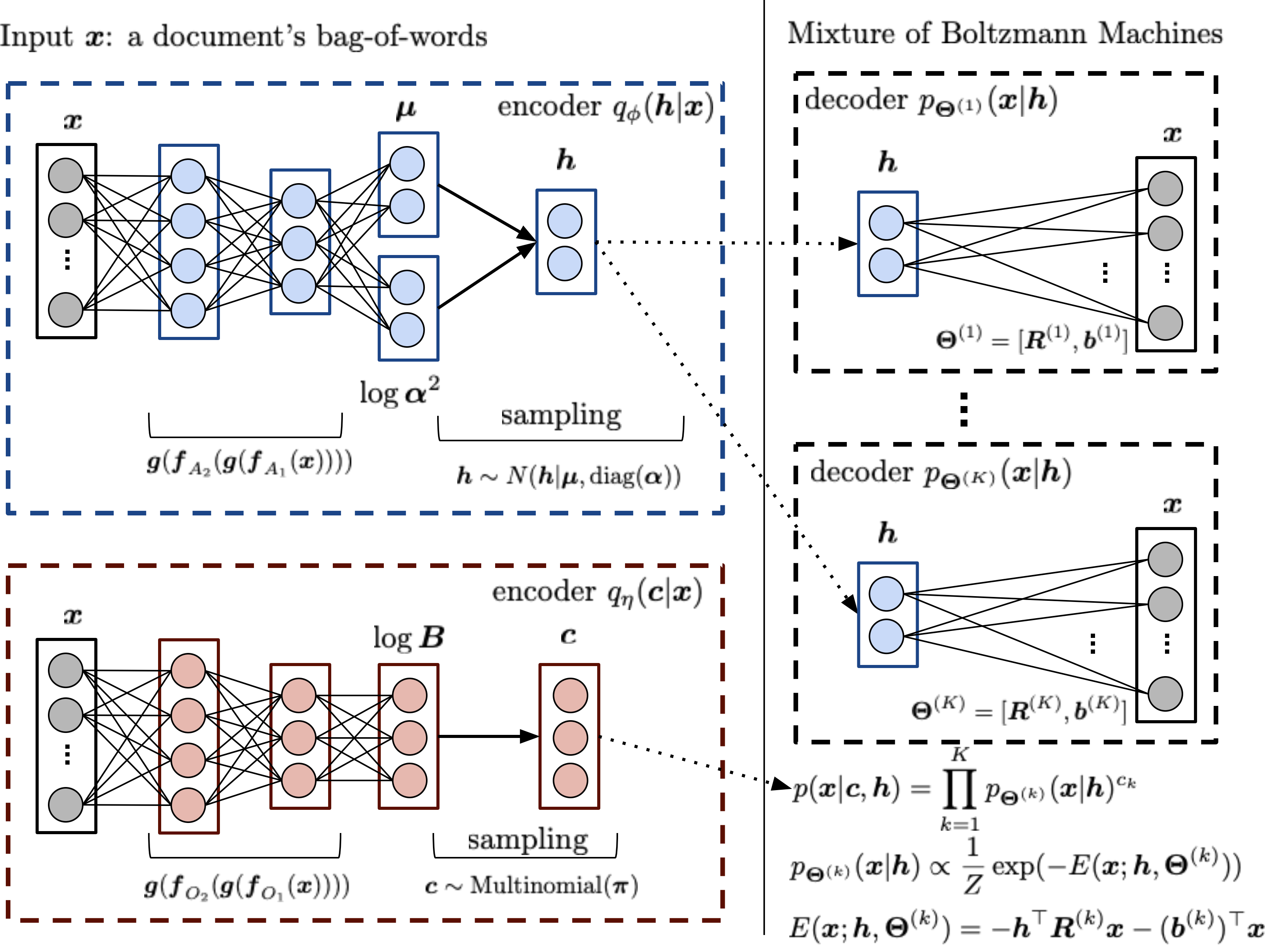

Input $\boldsymbol{x}$: a document's bag-of-words
Mixture of Boltzmann Machines
encoder $q_\phi(\boldsymbol{h}|\boldsymbol{x})$
$\boldsymbol{x}$
$\boldsymbol{\mu}$
$\boldsymbol{h}$
$\log \boldsymbol{\alpha}^2$
sampling
$\boldsymbol{g}(\boldsymbol{f}_{A_2}(\boldsymbol{g}(\boldsymbol{f}_{A_1}(\boldsymbol{x}))))$
$\boldsymbol{h} \sim N(\boldsymbol{h}|\boldsymbol{\mu}, \mathrm{diag}(\boldsymbol{\alpha}))$
decoder $p_{\boldsymbol{\Theta}^{(1)}}(\boldsymbol{x}|\boldsymbol{h})$
$\boldsymbol{\Theta}^{(1)} = [\boldsymbol{R}^{(1)}, \boldsymbol{b}^{(1)}]$
decoder $p_{\boldsymbol{\Theta}^{(K)}}(\boldsymbol{x}|\boldsymbol{h})$
$\boldsymbol{\Theta}^{(K)} = [\boldsymbol{R}^{(K)}, \boldsymbol{b}^{(K)}]$
encoder $q_\eta(\boldsymbol{c}|\boldsymbol{x})$
$\log \boldsymbol{B}$
$\boldsymbol{c}$
$\boldsymbol{g}(\boldsymbol{f}_{O_2}(\boldsymbol{g}(\boldsymbol{f}_{O_1}(\boldsymbol{x}))))$
$\boldsymbol{c} \sim \mathrm{Multinomial}(\boldsymbol{\pi})$
$p(\boldsymbol{x}|\boldsymbol{c}, \boldsymbol{h}) = \prod_{k=1}^{K} p_{\boldsymbol{\Theta}^{(k)}}(\boldsymbol{x}|\boldsymbol{h})^{c_k}$
$p_{\boldsymbol{\Theta}^{(k)}}(\boldsymbol{x}|\boldsymbol{h}) \propto \frac{1}{Z} \exp(-E(\boldsymbol{x}; \boldsymbol{h}, \boldsymbol{\Theta}^{(k)}))$
$E(\boldsymbol{x}; \boldsymbol{h}, \boldsymbol{\Theta}^{(k)}) = -\boldsymbol{h}^\top \boldsymbol{R}^{(k)} \boldsymbol{x} - (\boldsymbol{b}^{(k)})^\top \boldsymbol{x}$