(6.18) について考える。

(6.14) を考慮すると以下のようになる。

$$
\begin{aligned}
p(Y_U, Z_A, Z_U, W | X_A, X_U, Y_A) &= \frac{p(Y_U, Z_A, Z_U, W, X_A, X_U, Y_A)}{p(X_A, X_U, Y_A)} \\
&= \frac{p(X_A|Y_A, Z_A, W)p(Y_A)p(Z_A)p(X_U|Y_U, Z_U, W)p(Y_U)p(Z_U)p(W)}{p(X_A, X_U, Y_A)} \\
&= p(X_A|Y_A, Z_A, W)p(Z_A)p(X_U|Y_U, Z_U, W)p(Y_U)p(Z_U)p(W)\frac{p(Y_A)}{p(X_A, X_U, Y_A)} \\
&= p(X_A|Y_A, Z_A, W)p(Z_A)p(X_U|Y_U, Z_U, W)p(Y_U)p(Z_U)p(W)exp\{c\}
\end{aligned}
\tag{1}
$$

3 つめの等号のあとの分数の部分がデータとしてわかっているので定数。

また、q を平均場近似を用いて、以下のようにする。

$$
q \equiv q(Y_U, Z_A, Z_U, W; X_A, Y_A, X_U, \xi, \psi) \approx q(Y_U)q(Z_A)q(Z_U)q(W) \tag{2}
$$

これらは (6.6),(6.15)-(6.17) のように定数や、得られたデータを用いて、表される。

$$
\begin{aligned}
q &\equiv q(Y_U, Z_A, Z_U, W; X_A, Y_A, X_U, \xi, \psi) \\
&= q(Y_U; X_A, Y_A, X_U, \xi, \psi)q(Z_A; X_A, Y_A, X_U, \xi, \psi)q(Z_U; X_A, Y_A, X_U, \xi, \psi)q(W; X_A, Y_A, X_U, \xi, \psi) \\
&\approx q(Y_U)q(Z_A)q(Z_U)q(W) \approx q(Y_U; X_U, \psi)q(Z_A; X_A, Y_A, \psi)q(Z_U; X_U, \psi)q(W; \xi)
\end{aligned}
\tag{3}
$$

(3.10) の KL ダイバージェンスの定義を考慮すると、

$$
\begin{aligned}
&D_{KL}(q(Y_U, Z_A, Z_U, W; X_A, Y_A, X_U, \xi, \psi)||p(Y_U, Z_A, Z_U, W|X_A, X_U, Y_A)) = D_{KL}(q||p(Y_U, Z_A, Z_U, W|X_A, X_U, Y_A)) \\
&= -\int q\, ln\, \frac{p(Y_U, Z_A, Z_U, W|X_A, X_U, Y_A)}{q} dY_U dZ_A dZ_U dW \\
&= -\int q\, ln\, \frac{p(X_A|Y_A, Z_A, W)p(Z_A)p(X_U|Y_U, Z_U, W)p(Y_U)p(Z_U)p(W)exp\{c\}}{q(Y_U; X_U, \psi)q(Z_A; X_A, Y_A, \psi)q(Z_U; X_U, \psi)q(W; \xi)} dY_U dZ_A dZ_U dW \\
&= -\int q(Y_U; X_U, \psi)q(Z_U; X_U, \psi)dY_U dZ_U \int q(Z_A; X_A, Y_A, \psi)q(W; \xi)\, ln\, p(X_A|Y_A, Z_A, W)dZ_A dW \\
&\quad - \int q(Y_U; X_U, \psi)q(Z_U; X_U, \psi)q(W; \xi)dY_U dZ_U dW \int q(Z_A; X_A, Y_A, \psi)\, ln\, p(Z_A)dZ_A \\
&\quad + \int q(Y_U; X_U, \psi)q(Z_U; X_U, \psi)q(W; \xi)dY_U dZ_U dW \int q(Z_A; X_A, Y_A, \psi)\, ln\, q(Z_A; X_A, Y_A, \psi)dZ_A \\
&\quad - \int q(Z_A; X_A, Y_A, \psi)dZ_A \int q(Y_U; X_U, \psi)q(Z_U; X_U, \psi)q(W; \xi)\, ln\, p(X_U|Y_U, Z_U, W)dY_U dZ_U dW \\
&\quad - \int q(Z_A; X_A, Y_A, \psi)q(Z_U; X_U, \psi)q(W; \xi)dZ_A dZ_U dW \int q(Y_U; X_U, \psi)\, ln\, p(Y_U)dY_U \\
&\quad - \int q(Z_A; X_A, Y_A, \psi)q(Y_U; X_U, \psi)q(W; \xi)dZ_A dY_U dW \int q(Z_U; X_U, \psi)\, ln\, p(Z_U)dZ_U \\
&\quad + \int q(Z_U; X_U, \psi)q(W; \xi)q(Z_A; X_A, Y_A, \psi)dZ_A dZ_U dW \int q(Y_U; X_U, \psi)\, ln\, q(Y_U; X_U, \psi)dY_U \\
&\quad + \int q(Y_U; X_U, \psi)q(W; \xi)q(Z_A; X_A, Y_A, \psi)dZ_A dY_U dW \int q(Z_U; X_U, \psi)\, ln\, q(Z_U; X_U, \psi)dZ_U \\
&\quad - \int q(Z_A; X_A, Y_A, \psi)q(Y_U; X_U, \psi)q(Z_U; X_U, \psi)dZ_A dY_U dZ_U \int q(W; \xi)\, ln\, p(W)dW \\
&\quad + \int q(Y_U; X_U, \psi)q(Z_A; X_A, Y_A, \psi)q(Z_U; X_U, \psi)dZ_A dY_U dZ_U \int q(W; \xi)\, ln\, q(W; \xi)dW \\
&= -\int q(Z_A; X_A, Y_A, \psi)q(W; \xi)\, ln\, p(X_A|Y_A, Z_A, W)dZ_A dW - \int q(Z_A; X_A, Y_A, \psi)\, ln\, p(Z_A)dZ_A \\
&\quad + \int q(Z_A; X_A, Y_A, \psi)\, ln\, q(Z_A; X_A, Y_A, \psi)dZ_A - \int q(Y_U; X_U, \psi)q(Z_U; X_U, \psi)q(W; \xi)\, ln\, p(X_U|Y_U, Z_U, W)dY_U dZ_U dW \\
&\quad - \int q(Y_U; X_U, \psi)\, ln\, p(Y_U)dY_U - \int q(Z_U; X_U, \psi)\, ln\, p(Z_U)dZ_U \\
&\quad + \int q(Y_U; X_U, \psi)\, ln\, q(Y_U; X_U, \psi)dY_U + \int q(Z_U; X_U, \psi)\, ln\, q(Z_U; X_U, \psi)dZ_U - \int q(W; \xi)\, ln\, \frac{p(W)}{q(W; \xi)}dW + c \\
&= -\mathbb{E}_{q(Z_A; X_A, Y_A, \psi)q(W; \xi)}[ln\, p(X_A|Y_A, Z_A, W)] - \mathbb{E}_{q(Z_A; X_A, Y_A, \psi)}[ln\, p(Z_A)] + \mathbb{E}_{q(Z_A; X_A, Y_A, \psi)}[ln\, q(Z_A; X_A, Y_A, \psi)] \\
&\quad -\mathbb{E}_{q(Y_U; X_U, \psi)q(Z_U; X_U, \psi)q(W; \xi)}[ln\, p(X_U|Y_U, Z_U, W)] - \mathbb{E}_{q(Y_U; X_U, \psi)}[ln\, p(Y_U)] - \mathbb{E}_{q(Z_U; X_U, \psi)}[ln\, p(Z_U)] \\
&\quad +\mathbb{E}_{q(Y_U; X_U, \psi)}[ln\, q(Y_U; X_U, \psi)] + \mathbb{E}_{q(Z_U; X_U, \psi)}[ln\, q(Z_U; X_U, \psi)] - D_{KL}[q(W; \xi)||p(W)] + c
\end{aligned} \tag{4}
$$

これで (6.18) が求まった。ただし、全体を {} でくくっておらず、本の $\mathbb{E}_{q(Z_U; X_U, \psi)}[ln\, q(Z_U; X_U, \psi)]$ の符号は間違っているものと思われる。

さて、生成ネットワーク、推論ネットワークの変分パラメータはまとめて ψ としているが、これを、ψ_1, ψ_2 と分ける。すると、

$$
q(Y_U; X_U, \psi_2), q(Z_A; X_A, Y_A, \psi_1), q(Z_U; X_U, \psi_1) \tag{5}
$$

となる。すると

$$\frac{\partial \mathbb{F}}{\partial \psi_2} = \frac{\partial(-\mathbb{E}_{q(Y_U;X_U,\psi_2)q(Z_U;X_U,\psi_1)q(W;\xi)}[ln\,p(X_U|Y_U,Z_U,W)] - \mathbb{E}_{q(Y_U;X_U,\psi_2)}[ln\,p(Y_U)] + \mathbb{E}_{q(Y_U;X_U,\psi_2)}[ln\,q(Y_U;X_U,\psi_2)])}{\partial \psi_2} \tag{6}$$

となり、ψ_2 の更新に X_A, Y_A が含まれず、学習に反映されない。そのため、(6.22) にあるように、(6.18) の $-\mathbb{E}_{q(Y_U;X_U,\psi)}[ln\,q(Y_U;X_U,\psi)]$ に相当する項を追加して、X_A, Y_A を反映させて学習を行う。