1. Intro

Define variational free energy with a set of external auxiliary fields (Lagrange multipliers):

$$-\beta G = \ln \sum_{\mathbf{s}} \exp\left(\beta \sum_{(ij)} w_{ij} s_i s_j + \beta \sum_i a_i s_i + \sum_i \lambda_i(\beta)(s_i - m_i)\right)$$

We are interested in expanding $-\beta G(\beta, \mathbf{m})$ around $\beta = 0$ where the spins are entirely controlled by their auxiliary fields:

$$-\beta G = -(\beta G)_{\beta=0} - \left(\frac{\partial(\beta G)}{\partial \beta}\right)_{\beta=0} \beta - \left(\frac{\partial^2(\beta G)}{\partial \beta^2}\right)_{\beta=0} \frac{\beta^2}{2} - \ldots$$

At $\beta = 0$ the spins are uncorrelated (where $\mathbb{E}$ refers to the average configuration under the Boltzmann measure):

$$m_i = \mathbb{E}_{\beta=0}(s_i) = \frac{\exp(\lambda_i(0))}{\exp(\lambda_i(0)) + 1} = \text{sigmoid}(\lambda_i(0)) \tag{1}$$

and

$$\begin{aligned}
-(\beta G)_{\beta=0} &= \ln \sum_{\mathbf{s}} \exp\left(\sum_i \lambda_i(0)(s_i - m_i)\right) \\
&= \ln\left\{\sum_{s_1} \exp\left(\lambda_1(0)(s_1 - m_1)\right) \ldots \sum_{s_n} \exp\left(\lambda_n(0)(s_n - m_n)\right)\right\} \\
&= \ln\left\{(\exp(\lambda_1(0)) + 1)\exp(-\lambda_1(0)m_1) \ldots (\exp(\lambda_n(0)) + 1)\exp(-\lambda_n(0)m_n)\right\} \\
&= \sum_i \left(\ln\left[1 + \left(\frac{m_i}{1 - m_i}\right)\right] - \lambda_i(0) m_i\right)
\end{aligned} \tag{2}$$

Using 1, we obtain:

$$\lambda_i(0) = \text{logit}(m_i) = \ln\left(\frac{m_i}{1 - m_i}\right)$$

and thus:

$$\begin{aligned}
-(\beta G)_{\beta=0} &= \sum_i \left\{\ln\left(\frac{1}{1 - m_i}\right) - m_i \ln\left(\frac{m_i}{1 - m_i}\right)\right\} \\
&= -\sum_i \left[m_i \ln(m_i) + (1 - m_i)\ln(1 - m_i)\right]
\end{aligned}$$

Yedida and Georges showed how to create the Taylor expansion beyond $O(\beta^2)$ (TODO referring to MF and TAP variant) (derivation in Appendix):

$$\begin{aligned}
-\beta G = &-\sum_i \left[m_i \ln(m_i) + (1 - m_i)\ln(1 - m_i)\right] \\
&+ \beta \sum_{(ij)} w_{ij} m_i m_j + \beta \sum_i a_i m_i \\
&+ \frac{\beta^2}{2} \sum_{(ij)} w_{ij}^2 (m_i - m_i^2)(m_j - m_j^2) \\
&+ \frac{2\beta^3}{3} \sum_{(ij)} w_{ij}^3 (m_i - m_i^2)(\frac{1}{2} - m_i)(m_j - m_j^2)(\frac{1}{2} - m_j)
\end{aligned}$$

Appendix

Lets define energy as:

$$E = -\sum_{ij} w_{ij} s_i s_j - \sum_i a_i s_i \tag{3}$$

and introduce the following operator:

$$U \equiv E - \mathbb{E}(E) - \sum_i \frac{\partial \lambda_i(\beta)}{\partial \beta}(s_i - m_i) \tag{4}$$

which poses useful property:

$$\mathbb{E}(U) = 0$$

For any other operator O we then have:

$$\frac{\partial \mathbb{E}(O)}{\partial \beta} = \mathbb{E}\left(\frac{\partial O}{\partial \beta}\right) - \mathbb{E}(OU) \tag{5}$$

Now, the first derivative from the Taylor expansion is:

$$\begin{aligned}\frac{\partial(\beta G)}{\partial \beta} &= \frac{\sum_{\mathbf{s}} \exp\Big(\beta \sum_{(ij)} w_{ij} s_i s_j + \beta \sum_i a_i s_i + \sum_i \lambda_i(\beta)(s_i - m_i)\Big)\Big(\sum_{(ij)} w_{ij} s_i s_j + \sum_i a_i s_i + \sum_i \frac{\partial \lambda_i(\beta)}{\partial \beta}(s_i - m}{\sum_{\mathbf{s}} \exp\Big(\beta \sum_{(ij)} w_{ij} s_i s_j + \sum_i a_i s_i + \sum_i \lambda_i(\beta)(s_i - m_i)\Big)} \\ &= \sum_{(ij)} w_{ij} \mathbb{E}(s_i s_j) + \sum_i a_i \mathbb{E}(s_i) + \frac{\partial \lambda_i(\beta)}{\partial \beta} \sum_i \mathbb{E}(s_i - m_i)\end{aligned}$$

In the case of $\beta = 0$ we have:

$$\left.\frac{\partial(\beta G)}{\partial \beta}\right|_{\beta=0} = \sum_{(ij)} w_{ij} m_i m_j + \sum_i a_i m_i$$

Using 5 we obtain:

$$\frac{\partial m_i}{\partial \beta} = 0 = \frac{\partial \mathbb{E}(s_i)}{\partial \beta} = \mathbb{E}\left(\frac{\partial s_i}{\partial \beta}\right) - \mathbb{E}(s_i U) = -\mathbb{E}(s_i U) = -\mathbb{E}(U(s_i - m_i)) \tag{6}$$

Thus:

$$\begin{aligned}\frac{\partial U}{\partial \beta} &= \frac{\partial H}{\partial \beta} - \frac{\partial \mathbb{E}(H)}{\partial \beta} - \sum_i \frac{\partial^2 \lambda_i(\beta)}{\partial \beta^2}(s_i - m_i) \\ &= \mathbb{E}(U^2) - \sum_i \frac{\partial^2 \lambda_i(\beta)}{\partial \beta^2}(s_i - m_i)\end{aligned} \tag{7}$$

The second derivative has the form:

$$\begin{aligned}\frac{\partial^2 U}{\partial \beta^2} &= 2\mathbb{E}\left(\frac{\partial U}{\partial \beta} U\right) - \mathbb{E}(U^3) - \sum_i \frac{\partial^3 \lambda_i(\beta)}{\partial \beta^3}(s_i - m_i) \\ &= -\mathbb{E}(U^3) - \sum_i \frac{\partial^3 \lambda_i(\beta)}{\partial \beta^3}(s_i - m_i)\end{aligned} \tag{8}$$

Coming back to our expansion of free energy using formulas derived above we now can calculate:

$$\frac{\partial(\beta G)}{\partial \beta} = \mathbb{E}(E) - \sum_i \frac{\partial \lambda_i(\beta)}{\partial \beta} \mathbb{E}(s_i - m_i) = \mathbb{E}(E) \tag{9}$$

and:

$$\frac{\partial^2(\beta G)}{\partial \beta^2} = \mathbb{E}\left(\frac{\partial E}{\partial \beta}\right) - \mathbb{E}(EU) = -\mathbb{E}(U^2)$$

$$\frac{\partial^3(\beta G)}{\partial \beta^3} = -2\mathbb{E}\left(U\frac{\partial U}{\partial \beta}\right) + \mathbb{E}(U^3) = \mathbb{E}(U^3) \tag{10}$$

$$\frac{\partial^4(\beta G)}{\partial \beta^4} = 3\mathbb{E}\left(U^2\frac{\partial U}{\partial \beta}\right) - \mathbb{E}(U^4) = 3\left(\mathbb{E}(U^2)\right)^2 - 3\sum_i \frac{\partial^2 \lambda_i(\beta)}{\partial \beta^2}\mathbb{E}(U^2(s_i - m_i)) - \mathbb{E}(U^4)$$

TS was considered around point $\beta = 0$. Using derivations from above we obtain again a 'naive' term:

$$\left.\frac{\partial(\beta G)}{\partial \beta}\right|_{\beta=0} = \mathbb{E}_{\beta=0}(E) = -\sum_{(ij)} w_{ij}m_im_j - \sum_i a_im_i = -\frac{1}{2}\sum_i\sum_j w_{ij}m_im_j - \sum_i a_im_i \tag{11}$$

Consider now

$$\left.\frac{\partial(\beta G)}{\partial m_i \partial \beta}\right|_{\beta=0} = -\sum_{j\neq i} w_{ij}m_j - a_i \tag{12}$$

On the other hand:

$$\left.\frac{\partial(\beta G)}{\partial m_i \partial \beta}\right|_{\beta=0} = \left.\frac{\partial(\beta G)}{\partial \beta \partial m_i}\right|_{\beta=0} = \frac{\partial}{\partial \beta}\mathbb{E}(\lambda_i(\beta)) = \left.\frac{\partial \lambda_i(\beta)}{\partial \beta}\right|_{\beta=0} \tag{13}$$

Substituting 13 into 4 gives us:

$$\begin{aligned} U_{\beta=0} &= -\sum_{(ij)} w_{ij}s_is_j - \sum_i a_is_i + \frac{1}{2}\sum_i\sum_j w_{ij}m_im_j + \sum_i a_im_i + \sum_i\left(\sum_{j\neq i} w_{ij}m_j + a_i\right)(s_i - m_i) \\ &= -\sum_{(ij)} w_{ij}s_is_j - \frac{1}{2}\sum_{(ij)} w_{ij}m_im_j + \sum_i\sum_{j\neq i} w_{ij}s_im_j \\ &= -\sum_{(ij)} w_{ij}(s_i - m_i)(s_j - m_j) = -\sum_l w_ly_l \end{aligned} \tag{14}$$

where y_l stands for the 'link' operator $w_l = w_{ij}$ and $y_l = (s_i - m_i)(s_j - m_j)$ which poses useful properties:

$$\begin{aligned} \mathbb{E}(y_l)_{\beta=0} &= \mathbb{E}(s_is_j) - m_j\mathbb{E}(s_i) - m_i\mathbb{E}(s_j) + m_im_j = 0 \\ \mathbb{E}(y_l(s_i - m_i))_{\beta=0} &= m_j - m_j - m_i^2m_j + m_i^2m_j \\ &\quad - m_i^2m_j + m_i^2m_j + m_i^2m_j - m_i^2m_j \\ &= 0 \end{aligned} \tag{15}$$

Finally, if $k \neq l$ then:

$$\mathbb{E}(y_ky_l) = \mathbb{E}(y_k)\mathbb{E}(y_l) = 0$$

while for $k = l$ we have:

$$\begin{aligned} \mathbb{E}((s_i - m_1)^2(s_j - m_j)^2) &= m_im_j - 2m_im_j^2 + m_im_j^2 - 2m_i^2m_j + 4m_1^2m_j^2 \\ &\quad - 2m_i^2m_j^2 + m_i^2m_j - 2m_i^2m_j^2 + m_i^2m_j^2 \\ &= (m_i - m_i^2)(m_j - m_j^2) \end{aligned} \tag{16}$$

Using properties from 16 in equations 10 we can obtain:

$$\begin{aligned}\left.\frac{\partial^2(\beta G)}{\partial \beta^2}\right|_{\beta=0} &= -\ \mathbb{E}(U^2)_{\beta=0} \\ &= -\sum_{l_i l_2} w_{l_i} w_{l_2} \mathbb{E}_{\beta=0}(y_{l_1} y_{l_2}) \\ &= -\sum_{(i,j)} w_{ij}^2 (m_i - m_i^2)(m_j - m_j^2)\end{aligned}$$

which yields the TAP-Onsager term.

To obtain the next term for the Taylor expansion we need to compute $\mathbb{E}(y_{l_1} y_{l_2} y_{l_3})$ term and by definition the structure of the RBM model doesn't admit triangles in its corresponding factor graphs. Thus, we need to consider only the case when $l_1 = l_2 = l_3$:

$$\begin{aligned}\mathbb{E}((s_i - m_1)^3 (s_j - m_j)^3) = & \ m_i m_j - 3 m_i m_j^2 + 2 m_i m_j^2 + 2 m_i m_j^3 - 3 m_i^2 m_j + 2 m_i^3 m_j \\ & + 9 m_i^2 m_j^2 - 6 m_i^3 m_j^2 - 6 m_i^2 m_j^3 + 4 m_i^3 m_j^3 \\ = & \ 4(m_i - m_i^2)(\frac{1}{2} - m_i)(m_j - m_j^2)(\frac{1}{2} - m_j),\end{aligned} \tag{17}$$

which gives the third term:

$$\left.\frac{\partial^3(\beta G)}{\partial \beta^3}\right|_{\beta=0} = \frac{2\beta^3}{3} \sum_{(ij)} w_{ij}^3 (m_i - m_i^2)(\frac{1}{2} - m_i)(m_j - m_j^2)(\frac{1}{2} - m_j).$$