DeepMP derivation

October 24, 2021

1 Derivation of the BP and AMP equations

1.1 Derivation of the BP equations

In order to derive the BP equations, we start with the following portion of the factor graph reported in Eq. **??** in the main text, describing the contribution of a single data example in the inner loop of the PasP updates:

$$\prod_{\ell=0}^{L}\prod_{k} P^{\ell+1}\left(x_{kn}^{\ell+1}\;\middle|\;\sum_{i} W_{ki}^{\ell} x_{in}^{\ell}\right) \quad \text{where } \boldsymbol{x}_n^0=\boldsymbol{x}_n,\ \boldsymbol{x}_n^{L+1}=y_n. \tag{1}$$

where we recall that the quantity x_{kn}^{ℓ} corresponds to the activation of neuron k in layer ℓ in correspondence of the input example n.

Let us start by analyzing the single factor:

$$P^{\ell+1}\left(x_{kn}^{\ell+1}\;\middle|\;\sum_{i} W_{ki}^{\ell} x_{in}^{\ell}\right) \tag{2}$$

We refer to messages that travel from input to output in the factor graph as *upgoing* or *upwards* messages, while to the ones that travel from output to input as *downgoing* or *backwards* messages.

Factor-to-variable-W messages The factor-to-variable-W messages read:

$$\hat{\nu}_{kn\to ki}^{\ell+1}(W_{ki}^{\ell}) \propto \int \prod_{i'\neq i} d\nu_{ki'\to n}^{\ell}(W_{ki'}^{\ell}) \prod_{i'} d\nu_{i'n\to k}^{\ell}(x_{i'n}^{\ell})\; d\nu_{\downarrow}(x_{kn}^{\ell+1})\; P^{\ell+1}\left(x_{kn}^{\ell+1}\;\middle|\;\sum_{i'} W_{ki'}^{\ell} x_{i'n}^{\ell}\right) \tag{3}$$

where $\nu_{\downarrow}$ denotes the messages travelling downwards (from output to input) in the factor graph.

We denote the means and variances of the incoming messages respectively with $m^\ell_{ki\to n}$, $\hat{x}^\ell_{in\to k}$ and $\sigma^\ell_{ki\to n}$, $\Delta^\ell_{in\to k}$:

$$m^\ell_{ki\to n} = \int d\nu^\ell_{ki\to n}(W^\ell_{ki})\ W^\ell_{ki} \tag{4}$$

$$\sigma^\ell_{ki\to n} = \int d\nu^\ell_{ki\to n}(W^\ell_{ki})\ \left(W^\ell_{ki} - m^\ell_{ki\to n}\right)^2 \tag{5}$$

$$\hat{x}^\ell_{in\to k} = \int d\nu^\ell_{in\to k}(x^\ell_{in})\ x^\ell_{in} \tag{6}$$

$$\Delta^\ell_{in\to k} = \int d\nu^\ell_{in\to k}(x^\ell_{in})\ \left(x^\ell_{in} - \hat{x}^\ell_{in\to k}\right)^2 \tag{7}$$

We now use the central limit theorem to observe that with respect to the incoming messages distributions - assuming independence of these messages - in the large input limit the preactivation is a Gaussian random variable:

$$\sum_{i'\neq i} W^\ell_{ki'} x^\ell_{i'n} \sim \mathcal{N}(\omega^\ell_{kn\to i}, V^\ell_{kn\to i}) \tag{8}$$

where:

$$\omega^\ell_{kn\to i} = \mathbb{E}_\nu \left[\sum_{i'\neq i} W^\ell_{ki'} x^\ell_{i'n}\right] = \sum_{i'\neq i} m^\ell_{ki'\to n}\, \hat{x}^\ell_{i'n\to k} \tag{9}$$

$$\begin{aligned} V^\ell_{kn\to i} &= Var_\nu \left[\sum_{i'\neq i} W^\ell_{ki'} x^\ell_{i'n}\right] \\ &= \sum_{i'\neq i} \left(\sigma^\ell_{ki'\to n}\, \Delta^\ell_{i'n\to k} + \left(m^\ell_{ki'\to n}\right)^2\ \Delta^\ell_{i'n\to k} + \sigma^\ell_{ki'\to n}\ \left(\hat{x}^\ell_{i'n\to k}\right)^2\right) \end{aligned} \tag{10}$$

Therefore we can rewrite the outgoing messages as:

$$\hat{\nu}^{\ell+1}_{kn\to i}(W^\ell_{ki}) \propto \int dz\, d\nu^\ell_{in\to k}(x^\ell_{in})\, d\nu_\downarrow(x^{\ell+1}_{kn})\, e^{-\frac{\left(z-\omega_{kn\to i}-W^\ell_{ki}x^\ell_{in}\right)^2}{2V_{kn\to i}}}\ P^{\ell+1}\left(x^{\ell+1}_{kn}\ \middle|\ z\right) \tag{11}$$

We now assume $W^\ell_{ki} x^\ell_{in}$ to be small compared to the other terms. With a second order Taylor expansion we obtain:

$$\begin{aligned} \hat{\nu}^\ell_{kn\to i}(W^\ell_{ki}) \propto &\int dz\ \ d\nu_\downarrow(x^{\ell+1}_{kn})\ e^{-\frac{(z-\omega_{kn\to i})^2}{2V_{kn\to i}}}\, P^{\ell+1}\left(x^{\ell+1}_{kn}\ \middle|\ z\right) \\ &\times\left(1 + \frac{z-\omega_{kn\to i}}{V_{kn\to i}} \hat{x}^\ell_{in\to k} W^\ell_{ki} + \frac{\left(z-\omega_{kn\to i}\right)^2 - V_{kn\to i}}{2V_{kn\to i}}\left(\Delta + \left(\hat{x}^\ell_{in\to k}\right)^2\right)\left(W^\ell_{ki}\right)^2\right) \end{aligned} \tag{12}$$

Introducing now the function:

$$\varphi^\ell(B, A, \omega, V) = \log \int \mathrm{d}x\, \mathrm{d}z\ e^{-\frac{1}{2}Ax^2 + Bx}\ P^\ell\left(x|z\right) e^{-\frac{(\omega-z)^2}{2V}} \tag{13}$$

and defining:

$$g^{\ell}_{kn\to i} = \partial_\omega \varphi^{\ell+1}(B^{\ell+1}, A^{\ell+1}, \omega^{\ell}_{kn\to i}, V^{\ell}_{kn\to i}) \tag{14}$$
$$\Gamma^{\ell}_{kn\to i} = -\partial^2_\omega \varphi^{\ell+1}(B^{\ell+1}, A^{\ell+1}, \omega^{\ell}_{kn\to i}, V^{\ell}_{kn\to i}) \tag{15}$$

the expansion for the log-message reads:

$$\begin{aligned}\log \hat{\nu}^{\ell}_{kn\to i}(W^{\ell}_{ki}) \approx\, & const + \hat{x}^{\ell}_{in\to k}\, g^{\ell}_{kn\to i} W^{\ell}_{ki} \\ & - \frac{1}{2}\left(\left(\Delta^{\ell}_{in\to k} + \left(\hat{x}^{\ell}_{in\to k}\right)^2\right)\Gamma^{\ell}_{kn\to i} - \Delta^{\ell}_{in\to k}\left(g^{\ell}_{kn\to i}\right)^2\right)\left(W^{\ell}_{ki}\right)^2\end{aligned} \tag{16}$$

Factor-to-variable-x messages The derivation of these messages is analogous to the factor-to-variable-W ones in Eq. 3 just reported. The final result for the log-message is:

$$\begin{aligned}\log \hat{\nu}^{\ell}_{kn\to i}(x^{\ell}_{in}) \approx\, & const + m^{\ell}_{ki\to n}\, g^{\ell}_{kn\to i} x^{\ell}_{in} \\ & - \frac{1}{2}\left(\left(\sigma^{\ell}_{ki\to n} + \left(m^{\ell}_{ki\to n}\right)^2\right)\Gamma^{\ell}_{kn\to i} - \sigma^{\ell}_{ki\to n}\left(g^{\ell}_{kn\to i}\right)^2\right)\left(x^{\ell}_{in}\right)^2\end{aligned} \tag{17}$$

Variable-W-to-output factor messages The message from variable W^{ℓ}_{ki} to the output factor kn reads:

$$\begin{aligned}\nu^{\ell}_{ki\to n}(W^{\ell}_{ki}) &\propto P^{\ell}_{\theta_{ki}}(W^{\ell}_{ki}) e^{\sum_{n'\neq n} \log \hat{\nu}^{\ell}_{kn'\to i}(W^{\ell}_{ki})} \\ &\approx P^{\ell}_{\theta_{ki}}(W^{\ell}_{ki}) e^{H^{\ell}_{ki\to n} W^{\ell}_{ki} - \frac{1}{2} G^{\ell}_{ki\to n}\left(W^{\ell}_{ki}\right)^2}\end{aligned} \tag{18}$$

where we have defined:

$$H^{\ell}_{ki\to n} = \sum_{n'\neq n} \hat{x}^{\ell}_{in'\to k}\, g^{\ell}_{kn'\to i} \tag{19}$$
$$G^{\ell}_{ki\to n} = \sum_{n'\neq n} \left(\left(\Delta^{\ell}_{in'\to k} + \left(\hat{x}^{\ell}_{in'\to k}\right)^2\right)\Gamma^{\ell}_{kn'\to i} - \Delta^{\ell}_{in'\to k}\left(g^{\ell}_{kn'\to i}\right)^2\right) \tag{20}$$

Introducing now the effective free energy:

$$\psi^{\ell}(H, G, \theta) = \log \int \mathrm{d}W\;\; P^{\ell}_{\theta}(W)\, e^{HW - \frac{1}{2}GW^2} \tag{21}$$

we can express the first two cumulants of the message $\nu^{\ell}_{ki\to n}(W^{\ell}_{ki})$ as:

$$m^{\ell}_{ki\to n} = \partial_H \psi^{\ell}(H^{\ell}_{ki\to n}, G^{\ell}_{ki\to n}, \theta_{ki}) \tag{22}$$
$$\sigma^{\ell}_{ki\to n} = \partial^2_H \psi^{\ell}(H^{\ell}_{ki\to n}, G^{\ell}_{ki\to n}, \theta_{ki}) \tag{23}$$

Variable-x-to-input factor messages We can write the downgoing message as:

$$\begin{aligned}\nu_{\downarrow}(x_{in}^{\ell}) &\propto e^{\sum_k \log \hat{\nu}_{kn\to i}^{\ell}(x_{in}^{\ell})} \\ &\approx e^{B_{in}^{\ell} x - \frac{1}{2} A_{in}^{\ell} x^2} \end{aligned} \tag{24}$$

where:

$$B_{in}^{\ell} = \sum_n m_{ki\to n}^{\ell}\, g_{kn\to i}^{\ell} \tag{25}$$

$$A_{in}^{\ell} = \sum_n \left(\left(\sigma_{ki\to n}^{\ell} + \left(m_{ki\to n}^{\ell} \right)^2 \right) \Gamma_{kn\to i}^{\ell} - \sigma_{ki\to n}^{\ell} \left(g_{kn\to i}^{\ell+1} \right)^2 \right) \tag{26}$$

Variable-x-to-output factor messages By defining the following cavity quantities:

$$B_{in\to k}^{\ell} = B_{in\to k}^{\ell} - m_{ki\to n}^{\ell}\, g_{kn\to i}^{\ell} \tag{27}$$

$$A_{in\to k}^{\ell} = A_{in\to k}^{\ell} - \left(\left(\sigma_{ki\to n}^{\ell} + \left(m_{ki\to n}^{\ell} \right)^2 \right) \Gamma_{kn\to i}^{\ell} - \sigma_{ki\to n}^{\ell} \left(g_{kn\to i}^{\ell} \right)^2 \right) \tag{28}$$

and the following non-cavity ones:

$$\omega_{kn}^{\ell} = \sum_i m_{ki\to n}^{\ell}\, \hat{x}_{in\to k}^{\ell} \tag{29}$$

$$V_{kn}^{\ell} = \sum_i \left(\sigma_{ki\to n}^{\ell}\, \Delta_{in\to k}^{\ell} + \left(m_{ki\to n}^{\ell} \right)^2\, \Delta_{in\to k}^{\ell} + \sigma_{ki\to n}^{\ell}\, \left(\hat{x}_{i'n\to k}^{\ell} \right)^2 \right) \tag{30}$$

we can express the first 2 cumulants of the upgoing messages as:

$$\hat{x}_{in\to k}^{\ell} = \partial_B \varphi^{\ell}(B_{in\to k}^{\ell}, A_{in\to k}^{\ell}, \omega_{in}^{\ell-1}, V_{in}^{\ell-1}) \tag{31}$$

$$\Delta_{in\to k}^{\ell} = \partial_B^2 \varphi^{\ell}(B_{in\to k}^{\ell}, A_{in\to k}^{\ell}, \omega_{in}^{\ell-1}, V_{in}^{\ell-1}) \tag{32}$$

Wrapping it up Additional but straightforward considerations are required for the final input and output layers ($\ell = 0$ and $\ell = L + 1$ respectively), since they do not receive messages from below and above respectively. In the end, thanks to indipendence assumptions and the central limit theorem that we used throughout the derivations, we arrive to a closed set of equations involving the means and the variances (or otherwise the corresponding natural parameters) of the messages.

[TODO mention that we approximate cavity variances with non-cavity ones] [TODO use non-cavity variances for BP equations]

Dividing the update equations in a *forward* and *backward* pass, and ordering them using time indexes in such a way that we have an efficient flow of information, we obtain the set of BP equations presented in the main text Eqs. (**??**-**??**) and in the Appendix (**??**-**??**). We report the complete set of the BP equations here (TODO forse non c'Ãš bisogno, magari solo l'inizializzazione):

Initialization At $\tau = 0$:

$$B^{\ell,0}_{in\to k} = 0 \tag{33}$$
$$A^{\ell,0}_{in} = 0 \tag{34}$$
$$H^{\ell,0}_{ki\to n} = 0 \tag{35}$$
$$G^{\ell,0}_{ki} = 0 \tag{36}$$

Forward Pass At each $\tau = 1, \dots, \tau_{max}$, for $\ell = 0, \dots, L$ (TODO check ℓ range):

$$\hat{x}^{\ell,\tau}_{in\to k} = \partial_B \varphi^\ell(B^{\ell,\tau-1}_{in\to k}, A^{\ell,\tau-1}_{in}, \omega^{\ell-1,\tau}_{in}, V^{\ell-1,\tau}_{in}) \tag{37}$$
$$\Delta^{\ell,\tau}_{in} = \partial^2_B \varphi^\ell(B^{\ell,\tau-1}_{in}, A^{\ell,\tau-1}_{in}, \omega^{\ell-1,\tau}_{in}, V^{\ell-1,\tau}_{in}) \tag{38}$$
$$m^{\ell,\tau}_{ki\to n} = \partial_H \psi^\ell(H^{\ell,\tau-1}_{ki\to n}, G^{\ell,\tau-1}_{ki}, \theta^\ell_{ki}) \tag{39}$$
$$\sigma^{\ell,\tau}_{ki} = \partial^2_H \psi^\ell(H^{\ell,\tau-1}_{ki}, G^{\ell,\tau-1}_{ki}, \theta^\ell_{ki}) \tag{40}$$
$$V^{\ell,\tau}_{kn} = \sum_i \left(\left(m^{\ell,\tau}_{ki}\right)^2 \Delta^{\ell,\tau}_{in} + \sigma^{\ell,\tau-1}_{ki} \left(\hat{x}^{\ell,\tau}_{i'n}\right)^2 + \sigma^{\ell,\tau-1}_{ki} \Delta^{\ell,\tau}_{in} \right) \tag{41}$$
$$\omega^{\ell,\tau}_{kn\to i} = \sum_{i'\neq i} m^{\ell,\tau}_{ki'\to n}\, \hat{x}^{\ell,\tau}_{i'n\to k} \tag{42}$$

Backward Pass For $\tau = 1, \dots, \tau_{max}$, for $\ell = L, \dots, 0$ (TODO check ℓ range):

$$g^{\ell,\tau}_{kn\to i} = \partial_\omega \varphi^{\ell+1}(B^{\ell+1,\tau}_{kn}, A^{\ell+1,\tau}_{kn}, \omega^{\ell,\tau}_{kn\to i}, V^{\ell,\tau}_{kn}) \tag{43}$$
$$\Gamma^{\ell,\tau}_{kn} = -\partial^2_\omega \varphi^{\ell+1}(B^{\ell+1,\tau}_{kn}, A^{\ell+1,\tau}_{kn}, \omega^{\ell,\tau}_{kn}, V^{\ell,\tau}_{kn}) \tag{44}$$
$$A^{\ell,\tau}_{in} = \sum_k \left(\left(\left(m^{\ell,\tau}_{ki}\right)^2 + \sigma^{\ell,\tau}_{ki} \right) \Gamma^{\ell,\tau}_{kn} - \sigma^{\ell,\tau}_{ki} \left(g^{\ell,\tau}_{kn}\right)^2 \right) \tag{45}$$
$$B^{\ell,\tau}_{in\to k} = \sum_{k'\neq k} m^{\ell,\tau}_{k'i\to n}\, g^{\ell,\tau}_{k'n\to i} \tag{46}$$
$$G^{\ell,\tau}_{ki} = \sum_n \left(\left(\left(\hat{x}^{\ell,\tau}_{in}\right)^2 + \Delta^{\ell,\tau}_{in} \right) \Gamma^{\ell,\tau}_{kn} - \Delta^{\ell,\tau}_{in} \left(g^{\ell,\tau}_{kn}\right)^2 \right) \tag{47}$$
$$H^{\ell,\tau}_{ki\to n} = \sum_{n'\neq n} \hat{x}^{\ell,\tau}_{in'\to k}\, g^{\ell,\tau}_{kn'\to i} \tag{48}$$

1.2 Derivation of the AMP equations

In order to obtain the AMP equations, we approximate cavity quantities with non-cavity ones in the BP equations Eqs. (**??**-**??**) using a first order expansion.

We start with the mean activation:

$$
\begin{aligned}
\hat{x}^{\ell,\tau}_{in\to k} =& \partial_B \varphi^\ell(B^{\ell,\tau-1}_{in} - m^{\ell,\tau-1}_{ki\to n}\, g^{\ell,\tau-1}_{kn\to i}, A^{\ell,\tau-1}_{in}, \omega^{\ell-1,\tau}_{in}, V^{\ell-1,\tau}_{in}) \\
\approx& \partial_B \varphi^\ell(B^{\ell,\tau-1}_{in}, A^{\ell,\tau-1}_{in}, \omega^{\ell-1,\tau}_{in}, V^{\ell-1,\tau}_{in}) \\
&- m^{\ell,\tau-1}_{ki\to n}\, g^{\ell,\tau-1}_{kn\to i} \partial^2_B \varphi^\ell(B^{\ell,\tau-1}_{in}, A^{\ell,\tau-1}_{in}, \omega^{\ell-1,\tau}_{in}, V^{\ell-1,\tau}_{in}) \\
\approx& \hat{x}^{\ell,\tau}_{in} - m^{\ell,\tau-1}_{ki} g^{\ell,\tau-1}_{kn} \Delta^{\ell,\tau}_{in}
\end{aligned} \tag{49}
$$

Analogously, for the weight's mean we have:

$$
\begin{aligned}
m^{\ell,\tau}_{ki\to n} &= \partial_H \psi^\ell(H^{\ell,\tau-1}_{ki} - \hat{x}^{\ell,\tau-1}_{in\to k}\, g^{\ell,\tau-1}_{kn\to i}, G^{\ell,\tau-1}_{ki}, \theta^\ell_{ki}) \\
&\approx \partial_H \psi^\ell(H^{\ell,\tau-1}_{ki}, G^{\ell,\tau-1}_{ki}, \theta^\ell_{ki}) - \hat{x}^{\ell,\tau-1}_{in\to k}\, g^{\ell,\tau-1}_{kn\to i} \partial^2_H \psi^\ell(H^{\ell,\tau-1}_{ki}, G^{\ell,\tau-1}_{ki}, \theta^\ell_{ki}) \\
&\approx m^{\ell,\tau}_{ki} - \hat{x}^{\ell,\tau-1}_{in}\, g^{\ell,\tau-1}_{kn}\, \sigma^{\ell,\tau}_{ki}.
\end{aligned} \tag{50}
$$

This brings us to:

$$
\begin{aligned}
\omega^{\ell,\tau}_{kn} &= \sum_i m^{\ell,\tau}_{ki\to n}\, \hat{x}^{\ell,\tau}_{in\to k} \\
&\approx \sum_i m^{\ell,\tau}_{ki}\, \hat{x}^{\ell,\tau}_{in} - g^{\ell,\tau-1}_{kn} \sum_i \sigma^{\ell,\tau}_{ki} \hat{x}^{\ell,\tau}_{in} \hat{x}^{\ell,\tau-1}_{in} - g^{\ell,\tau-1}_{kn} \sum_i m^{\ell,\tau}_{ki} m^{\ell,\tau-1}_{ki} \Delta^{\ell,\tau}_{in} \\
&\quad [maybe - irrelevant] + (g^{\ell,\tau-1}_{kn})^2 \sum_i \sigma^{\ell,\tau}_{ki} m^{\ell,\tau-1}_{ki} \hat{x}^{\ell,\tau-1}_{in} \Delta^{\ell,\tau}_{in}
\end{aligned} \tag{51}
$$

Let us now apply the same procedure to the other set of cavity messages:

$$
\begin{aligned}
g^{\ell,\tau}_{kn\to i} =& \partial_\omega \varphi^{\ell+1}(B^{\ell+1,\tau}_{kn}, A^{\ell+1,\tau}_{kn}, \omega^{\ell,\tau}_{kn} - m^{\ell,\tau}_{ki\to n}\, \hat{x}^{\ell,\tau}_{in\to k}, V^{\ell,\tau}_{kn}) \\
\approx& \partial_\omega \varphi^{\ell+1}(B^{\ell+1,\tau}_{kn}, A^{\ell+1,\tau}_{kn}, \omega^{\ell,\tau}_{kn}, V^{\ell,\tau}_{kn}) \\
&- m^{\ell,\tau}_{ki\to n}\, \hat{x}^{\ell,\tau}_{in\to k} \partial^2_\omega \varphi^{\ell+1}(B^{\ell+1,\tau}_{kn}, A^{\ell+1,\tau}_{kn}, \omega^{\ell,\tau}_{kn}, V^{\ell,\tau}_{kn}) \\
\approx& g^{\ell,\tau}_{kn} + m^{\ell,\tau}_{ki} \hat{x}^{\ell,\tau}_{in} \Gamma^{\ell,\tau}_{kn}
\end{aligned} \tag{52}
$$

$$
\begin{aligned}
B^{\ell,\tau}_{in} &= \sum_k m^{\ell,\tau}_{ki\to n}\, g^{\ell,\tau}_{kn\to i} \\
&\approx \sum_k m^{\ell,\tau}_{ki}\, g^{\ell,\tau}_{kn} - \hat{x}_{in} \sum_k \left(g^{\ell,\tau}_{kn}\right)^2 \sigma^{\ell,\tau}_{ki} + \hat{x}^{\ell,\tau}_{in} \sum_k (m^{\ell,\tau}_{ki})^2 \Gamma^{\ell,\tau}_{kn} \\
&\quad [maybe - irrelevant] - (\hat{x}^{\ell,\tau}_{in})^2 \sum_k \sigma^{\ell,\tau}_{ki} m^{\ell,\tau}_{ki} g^{\ell,\tau}_{kn} \Gamma^{\ell,\tau}_{kn}
\end{aligned} \tag{53}
$$

$$
\begin{aligned}
H^{\ell,\tau}_{ki} &= \sum_n \hat{x}^{\ell,\tau}_{in\to k}\, g^{\ell,\tau}_{kn\to i} \\
&\approx \sum_n \hat{x}^{\ell,\tau}_{in}\, g^{\ell,\tau}_{kn} + m^{\ell,\tau}_{ki} \sum_n \left(\hat{x}^{\ell,\tau}_{in}\right)^2 \Gamma^{\ell,\tau}_{kn} - m^{\ell,\tau}_{ki} \sum_n (g^{\ell,\tau}_{kn})^2 \Delta^{\ell,\tau}_{in} \\
&\quad [maybe - irrelevant] - (m^{\ell,\tau}_{ki})^2 \sum_n g^{\ell,\tau}_{kn} \Gamma^{\ell,\tau}_{kn} \Delta^{\ell,\tau}_{in} \hat{x}^{\ell,\tau}_{in}
\end{aligned} \tag{54}
$$

We are now able to write down the full AMP equations:

Initialization At $\tau = 0$:

$$B_{in}^{\ell,0} = 0 \tag{55}$$
$$A_{in}^{\ell,0} = 0 \tag{56}$$
$$H_{ki}^{\ell,0} = 0 \text{ or some values} \tag{57}$$
$$G_{ki}^{\ell,0} = 0 \text{ or some values} \tag{58}$$
$$g_{kn}^{\ell,0} = 0 \tag{59}$$

Forward Pass At each $\tau = 1, \ldots, \tau_{max}$, for $\ell = 0, \ldots, L$ (TODO check ℓ range):

$$\hat{x}_{in}^{\ell,\tau} = \partial_B \varphi^\ell(B_{in}^{\ell,\tau-1}, A_{in}^{\ell,\tau-1}, \omega_{in}^{\ell-1,\tau}, V_{in}^{\ell-1,\tau}) \tag{60}$$
$$\Delta_{in}^{\ell,\tau} = \partial_B^2 \varphi^\ell(B_{in}^{\ell,\tau-1}, A_{in}^{\ell,\tau-1}, \omega_{in}^{\ell-1,\tau}, V_{in}^{\ell-1,\tau}) \tag{61}$$
$$m_{ki}^{\ell,\tau} = \partial_H \psi^\ell(H_{ki}^{\ell,\tau-1}, G_{ki}^{\ell,\tau-1}, \theta_{ki}^\ell) \tag{62}$$
$$\sigma_{ki}^{\ell,\tau} = \partial_H^2 \psi^\ell(H_{ki}^{\ell,\tau-1}, G_{ki}^{\ell,\tau-1}, \theta_{ki}^\ell) \tag{63}$$
$$V_{kn}^{\ell,\tau} = \sum_i \left(\left(m_{ki}^{\ell,\tau} \right)^2 \Delta_{in}^{\ell,\tau} + \sigma_{ki}^{\ell,\tau} \left(\hat{x}_{i'n}^{\ell,\tau} \right)^2 + \sigma_{ki}^{\ell,\tau} \Delta_{in}^{\ell,\tau} \right) \tag{64}$$
$$\begin{aligned}\omega_{kn}^{\ell,\tau} = &\sum_i m_{ki}^{\ell,\tau} \hat{x}_{in}^{\ell,\tau} - g_{kn}^{\ell,\tau-1} \sum_i \sigma_{ki}^{\ell,\tau} \hat{x}_{in}^{\ell,\tau} \hat{x}_{in}^{\ell,\tau-1} - g_{kn}^{\ell,\tau-1} \sum_i m_{ki}^{\ell,\tau} m_{ki}^{\ell,\tau-1} \Delta_{in}^{\ell,\tau} \\ &+ (g_{kn}^{\ell,\tau-1})^2 \sum_i \sigma_{ki}^{\ell,\tau} m_{ki}^{\ell,\tau-1} \hat{x}_{in}^{\ell,\tau-1} \Delta_{in}^{\ell,\tau}\end{aligned} \tag{65}$$

Backward Pass

$$g_{kn}^{\ell,\tau} = \partial_\omega \varphi^{\ell+1}(B_{kn}^{\ell+1,\tau}, A_{kn}^{\ell+1,\tau}, \omega_{kn\to i}^{\ell,\tau}, V_{kn}^{\ell,\tau}) \tag{66}$$
$$\Gamma_{kn}^{\ell,\tau} = -\partial_\omega^2 \varphi^{\ell+1}(B_{kn}^{\ell+1,\tau}, A_{kn}^{\ell+1,\tau}, \omega_{kn}^{\ell,\tau}, V_{kn}^{\ell,\tau}) \tag{67}$$
$$A_{in}^{\ell,\tau} = \sum_k \left(\left(\left(m_{ki}^{\ell,\tau} \right)^2 + \sigma_{ki}^{\ell,\tau} \right) \Gamma_{kn}^{\ell,\tau} - \sigma_{ki}^{\ell,\tau} \left(g_{kn}^{\ell,\tau} \right)^2 \right) \tag{68}$$
$$\begin{aligned}B_{in}^{\ell,\tau} = &\sum_k m_{ki}^{\ell,\tau} g_{kn}^{\ell,\tau} - \hat{x}_{in} \sum_k \left(g_{kn}^{\ell,\tau} \right)^2 \sigma_{ki}^{\ell,\tau} + \hat{x}_{in}^{\ell,\tau} \sum_k (m_{ki}^{\ell,\tau})^2 \Gamma_{kn}^{\ell,\tau} \\ &- (\hat{x}_{in}^{\ell,\tau})^2 \sum_k \sigma_{ki}^{\ell,\tau} m_{ki}^{\ell,\tau} g_{kn}^{\ell,\tau} \Gamma_{kn}^{\ell,\tau}\end{aligned} \tag{69}$$
$$G_{ki}^{\ell,\tau} = \sum_n \left(\left(\left(\hat{x}_{in}^{\ell,\tau} \right)^2 + \Delta_{in}^{\ell,\tau} \right) \Gamma_{kn}^{\ell,\tau} - \Delta_{in}^{\ell,\tau} \left(g_{kn}^{\ell,\tau} \right)^2 \right) \tag{70}$$
$$\begin{aligned}H_{ki}^{\ell,\tau} = &\sum_n \hat{x}_{in}^{\ell,\tau} g_{kn}^{\ell,\tau} + m_{ki}^{\ell,\tau} \sum_n \left(\hat{x}_{in}^{\ell,\tau} \right)^2 \Gamma_{kn}^{\ell,\tau} - m_{ki}^{\ell,\tau} \sum_n (g_{kn}^{\ell,\tau})^2 \Delta_{in}^{\ell,\tau} \\ &- (m_{ki}^{\ell,\tau})^2 \sum_n g_{kn}^{\ell,\tau} \Gamma_{kn}^{\ell,\tau} \Delta_{in}^{\ell,\tau} \hat{x}_{in}^{\ell,\tau}\end{aligned} \tag{71}$$

TODO

- B and H equations have to be changed yet on the cl/amp branch in DeepMP.jl

- not clear how to put damping

- presentare in appendice una sola volta la forma finale delle equazioni di BP e AMP

- nei vari algoritmi riportati nel paper non viene specificata l'inizializzazione dei messaggi