Tutorial: Approximate Message Passing & Replicas

Benjamin. Aubin, Marylou Gabrié
Institut de Physique Theorique, CEA-Saclay and
Ecole Normale Supérieure, Paris
(Dated: February 13, 2019)

Contents

I. General setting

In these notes, we provide a short introduction to Approximate Message Passing and Replicas computation in a Generalized Linear Model (GLM).

References

For more details see [1], [2], [3], [4], [5], [?]

A. Notations

- $\underline{\underline{X}} \in \mathbb{R}^{N\times M}$ contains the data as M $N-$dimensional samples, $i.i.d$ distributed $P_x(\underline{\underline{X}}) = \prod_{i,\mu=1}^{N,M} P_x(X_{ij}) \sim \mathcal{N}(0,1)$.
- $\underline{w} \in \mathbb{R}^{1\times N}$ is the matrix of weights of the second layer with prior: $P_w(\underline{w}) = \prod_{i=1}^{N} P_w(\underline{w}_i)$
- $\underline{\underline{y}} \in \mathbb{R}^{1\times M}$ is a set of M scalar observations.
- ϕ denotes activation functions
- Indices $\mu \in [1:M]$ and $i \in [1:N]$ correspond respectively to data samples and variables.
- $\underline{\underline{\eta}} \in \mathbb{R}^{1\times M}$ is a noise matrix applied to the second layer $\underline{\underline{\eta}} \sim \mathcal{N}(\underline{0}, \underline{\underline{\Delta}})$, with $\underline{\underline{\Delta}} = \Delta\underline{\underline{\mathbb{1}}}$
- $\alpha \equiv \frac{M}{N}$ with $N, M \to \infty$, $\alpha = \mathcal{O}(1)$

B. Model

a. Teacher-student scenario

We revisit the teacher-student average scenario: a teacher generates a training set using a *planted* solution $\underline{w}^0$. Then the student tries to learn/infer the teacher solution, using the training set generated by the teacher.

- *Teacher*

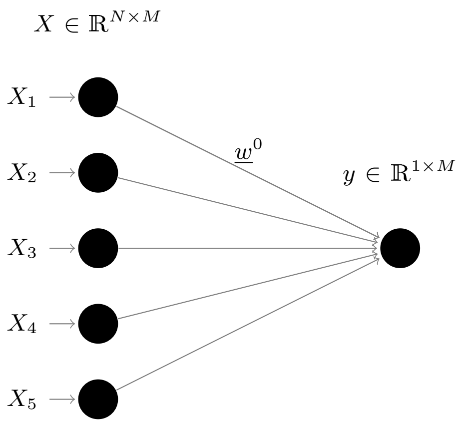
1. Data $\{\underline{X}_\mu\}_{\mu=1}^{M}$ are drawn *iid* (along both axis): $X_{ij} \sim \mathcal{N}(0,1)$
2. The teacher draws "planted" weights $\underline{w}^0$ from P_{w^0}
3. Finally, he generates a data set $\{y_\mu, \underline{X}_\mu\}_{\mu=1}^{M}$ using its activation function ϕ^0 and a gaussian noise $\eta^0 \sim \mathcal{N}(0, \Delta^0)$, according to:

$$y_\mu = \left(\phi^0\left(\sum_{i=1}^{N} \frac{1}{\sqrt{N}} w_i^0 X_{\mu i}\right) + \eta_\mu^0\right) \iff \underline{y} = \phi^0\left(\frac{1}{\sqrt{N}}\underline{w}^0\underline{\underline{X}}\right) + \underline{\eta}^0 \equiv \varphi_{out}^0\left(\frac{1}{\sqrt{N}}\underline{w}^0\underline{\underline{X}}, \Delta^0\right)$$

We define the teacher channel distribution, that will be involved later on:

$$P_{out}\left(y_\mu|\varphi_{out}^0\left(\frac{1}{\sqrt{N}}\underline{w}^0\underline{X}_\mu;\Delta^0\right)\right) \equiv \int dP(\underline{\eta}) P_{out}^0\left(y_\mu|\varphi_{out}\left(\frac{1}{\sqrt{N}}\underline{wX}_\mu;\underline{\eta}^0\right)\right) \tag{1}$$

- *Student*



1. The student tries to learn $\underline{w}^0$ from the dataset $\{y_\mu, \underline{X}_\mu\}_{\mu=1}^{M}$
2. We consider that the student has the same architecture with activation function ϕ, and has a prior distribution P_w on the weights $\underline{w}$ according to:

$$y_\mu = \left(\phi\left(\sum_{i=1}^{N}\frac{1}{\sqrt{N}}w_i X_{\mu i}\right) + \eta_\mu\right) \iff \underline{y} = \phi\left(\frac{1}{\sqrt{N}}\underline{w}\underline{\underline{X}}\right) + \underline{\eta} \equiv \varphi_{out}\left(\frac{1}{\sqrt{N}}\underline{w}\underline{\underline{X}}, \Delta\right)$$

As above, we define the student channel distribution:

$$P_{out}\left(y_\mu|\varphi_{out}\left(\frac{1}{\sqrt{N}}\underline{w}\underline{X}_\mu;\Delta\right)\right) \equiv \int dP(\underline{\eta})P_{out}\left(y_\mu|\varphi_{out}\left(\frac{1}{\sqrt{N}}\underline{w}\underline{X}_\mu;\underline{\eta}\right)\right) \tag{2}$$

- *Example*

For the gaussian channel:

$$\begin{cases} \phi(z) = z \\ \varphi_{out}\left(z,\Delta\right) = \phi(z) + \sqrt{\Delta}\xi = z + \sqrt{\Delta}\xi \\ P_{out}\left(y|\varphi_{out}\left(z;\Delta^0\right)\right) = \frac{e^{-\frac{1}{2\Delta}(y-z)^2}}{\sqrt{2\pi\Delta}} \end{cases}$$

C. Bayesian inference and factor graph

a. Problem statement

Our goal is to estimate the high dimensional ($N, M \to \infty$) probability distribution over $P\left(\underline{w}|\underline{y};\underline{\underline{X}}\right)$. Assuming that prior and channel distributions factorize (in one dimension) and using the Bayes formula, we can write it as:

$$P\left(\underline{w}|\underline{y};\underline{\underline{X}}\right) = \frac{P\left(\underline{y}|\underline{w};\underline{\underline{X}}\right)P\left(\underline{w}\right)}{P\left(\underline{y};\underline{\underline{X}}\right)} = \frac{P_{out}\left(\underline{y}|\varphi_{out}\left(\frac{1}{\sqrt{N}}\underline{w}\underline{\underline{X}};\Delta\right)\right)P_w(\underline{w})}{P(\underline{y};\underline{\underline{X}})} \tag{3}$$

$$= \frac{1}{P(\underline{y};\underline{\underline{X}})}\prod_{i=1}^{N}P_w\left(w_i\right)\prod_{\mu=1}^{M}P_{out}\left(\underline{y}_\mu|\varphi_{out}\left(\frac{1}{\sqrt{N}}\underline{w}\underline{X}_\mu;\Delta\right)\right) \tag{4}$$

Note that $\mathcal{Z}\left(\underline{y};\underline{\underline{X}}\right) \equiv P(\underline{y};\underline{\underline{X}}) \equiv$ plays the role of the partition function. Also, this distribution can be represented by the following factor graph:

COMMENT:

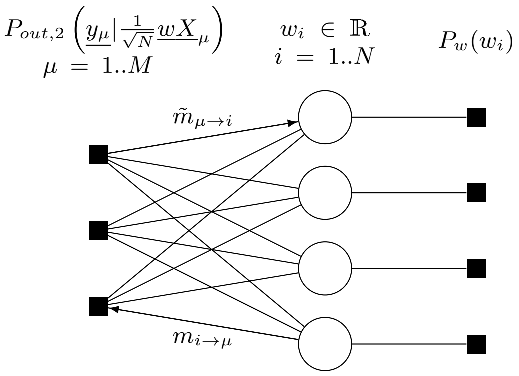


Figure 1: Factor graph for $M = 3$, $N = 4$

- The distribution $P\left(\underline{w}|\underline{y};\underline{\underline{X}}\right)$ is often intractable
- Hard to sample efficiently

COMMENT: Instead

- We may focus only on marginals $P(w_i)$ that can been estimated with Approximate Message Passing algorithms: Part.II
- We may also try to compute the free entropy $\Phi = \mathbb{E}_{\underline{y},\underline{w}^0,\underline{\underline{X}}}\left[\log\left(\mathcal{Z}\left(\underline{y};\underline{\underline{X}}\right)\right)\right]$: Part.III
- We try to give an idea how these two methods relate, and show they are consistent.

II. Approximate Message Passing algorithm

As we stressed above, the aim is to estimate the marginal probabilities of the variable w_i. We present here a short overview how to derive the AMP algorithm. The idea is to iterate partial "beliefs" between nodes of the graph until convergence.

A. Step1: Belief Propagation (BP) equations

Consider the factor graph 1. The first step is to write down the belief propagation equations. To do so, we add some *messages/beliefs* $m_{i\to\mu}(w_i)$ and $\tilde{m}_{\mu\to i}(w_i)$ on the edges of the factor graph. Each message is the marginal probability of the variable w_i if we remove the edge between variable i and constraint μ. BP equations read as follows:

$$\begin{cases} m_{i\to\mu}^{t+1}(w_i) = \dfrac{1}{\mathcal{Z}_{i\to\mu}} P_0(w_i) \displaystyle\prod_{k\neq\mu}^{M} \tilde{m}_{\nu\to i}^{t}(w_i) \\ \tilde{m}_{\mu\to i}^{t}(w_i) = \dfrac{1}{\mathcal{Z}_{\mu\to i}} \displaystyle\int \prod_{j\neq i}^{N} dw_j P_{\text{out}}\left(y_\mu | \frac{1}{\sqrt{N}} \sum_{j=1}^{N} X_{\mu j} w_j\right) m_{j\to\mu}^{t}(w_j)\,, \end{cases} \tag{5}$$

where we assumed that messages are independent. Hence these equations are exact on a tree (no loop), but they remain exact if correlations decrease fast enough / long loops.

The idea is to close these equations over the mean and variance of the variables (see Gaussian hypothesis / expansion in $\mathcal{O}(N)$ in IV A). We define then the estimated mean and variance at time t of the variable j.

$$\begin{cases} \hat{w}_{j\to\mu}^{t} \equiv \int_{\mathbb{R}} dw_j m_{j\to\mu}^{t}(w_j) w_j \\ \hat{c}_{j\to\mu}^{t} \equiv \int_{\mathbb{R}} dw_j m_{j\to\mu}^{t}(w_j) w_j w_j^{\intercal} - \hat{w}_{j\to\mu}^{t} (\hat{w}_{j\to\mu}^{t})^{\intercal} \end{cases}$$

B. Step 2: Towards relaxed BP

The idea is to close equations over mean and variance $\hat{w}_{j\to\mu}^{t}$ and $\hat{c}_{j\to\mu}^{t}$ using:

1. Fourier transform to decouple variables
2. $N \to \infty$ expansion

We show the derivation in IV A. In the end we obtain a set of $\mathcal{O}(N^2)$ messages, called Relaxed BP equations.

Summary of the Relaxed BP set of equations

In the end, Relaxed BP equations are simply the following set of equations:

$$\begin{cases} \hat{w}_{i\to\mu}^{t+1} = f_1^w(T_{\mu\to i}^t, \Sigma_{\mu\to i}^t) \\ \hat{c}_{i\to\mu}^{t+1} = f_2^w(T_{\mu\to i}^t, \Sigma_{\mu\to i}^t) \\ \Sigma_{\mu\to i}^t = \left(\sum_{\nu\neq\mu}^{M} A_{\nu\to i}^t\right)^{-1} \\ T_{\mu\to i}^t = \Sigma_{\mu\to i}^t \left(\sum_{\nu\neq\mu}^{M} B_{\nu\to i}^t\right) \end{cases} \qquad \begin{cases} B_{\mu\to i}^t = \frac{X_{\mu i}}{\sqrt{N}} g_{\text{out}}(\omega_{i\mu}^t, y_\mu, V_{i\mu}^t) \\ A_{\mu\to i}^t = -\frac{X_{\mu i}^2}{N} \partial_\omega g_{\text{out}}(\omega_{i\mu}^t, y_\mu, V_{i\mu}^t) \\ \omega_{i\mu}^t = \sum_{j\neq i}^{N} \frac{X_{\mu j}}{\sqrt{N}} \hat{w}_{j\to\mu}^t \\ V_{i\mu}^t = \sum_{j\neq i}^{N} \frac{X_{\mu j}^2}{N} \hat{c}_{j\to\mu}^t \end{cases} \tag{6}$$

where f_1^w and f_2^w are defined in II E.

C. Step 3: Towards Approximate Message Passing algorithm

The relaxed BP algorithm uses $\mathcal{O}(N^2)$ messages. However all the messages depend weakly on the target node. The missing message is negligible, that allows us to expand the previous Relaxed BP equations (II B) in the $N \to \infty$ limit.

We define the following messages, where we removed the target node dependence:

$$\begin{cases} \omega_\mu^t \equiv \sum_{j=1}^{N} \frac{X_{\mu j}}{\sqrt{N}} \hat{w}_{j\to\mu}^t \\ V_\mu^t \equiv \sum_{j=1}^{N} \frac{X_{\mu j}^2}{N} \hat{c}_{j\to\mu}^t \end{cases} \qquad \begin{cases} \Sigma_i^t \equiv \left(\sum_{\nu=1}^{M} A_{\nu\to i}^t \right)^{-1} \\ T_i^t \equiv \Sigma_i^t \left(\sum_{\nu=1}^{M} B_{\nu\to i}^t \right) \end{cases} \tag{7}$$

Expanding (II B) around the above "full" messages, we obtain a closed set of equation over only $\mathcal{O}(N)$ variables, called Generalized Approximate Message Passing (GAMP). We present the derivation in (Eq.IV B). Finally the algorithm reads:

Summary - AMP algorithm

Input: vector $y \in \mathbb{R}^M$ and matrix $X \in \mathbb{R}^{M\times N}$:
Initialize: $\hat{w}_i$, $g_{\text{out},\mu} \in \mathbb{R}$ and $\hat{c}_i$, $\partial_\omega g_{\text{out},\mu} \in \mathbb{R}^+$ for $1 \le i \le N$ and $1 \le \mu \le M$ at $t = 0$.
repeat

Update of the mean $\omega_\mu \in \mathbb{R}$ and covariance $V_\mu \in \mathbb{R}^+$:

$$\omega_\mu^t = \sum_{i=1}^{N} \left(\frac{X_{\mu i}}{\sqrt{N}} \hat{w}_i^t - \frac{X_{\mu i}^2}{N} \left(\Sigma_i^{t-1}\right)^{-1} \hat{c}_i^t \Sigma_i^{t-1} g_{\text{out},\mu}^{t-1} \right)$$

$$V_\mu^t = \sum_{i=1}^{N} \frac{X_{\mu i}^2}{N} \hat{c}_i^t$$

Update of $g_{\text{out},\mu} \in \mathbb{R}$ and $\partial_\omega g_{\text{out},\mu} \in \mathbb{R}^+$:

$$g_{\text{out},\mu}^t = g_{\text{out}}(\omega_\mu^t, Y_\mu, V_\mu^t)$$

$$\partial_\omega g_{\text{out},\mu}^t = \partial_\omega g_{\text{out}}(\omega_\mu^t, Y_\mu, V_\mu^t)$$

Update of the mean $T_i \in \mathbb{R}$ and covariance $\Sigma_i \in \mathbb{R}^+$:

$$T_i^t = \Sigma_i^t \left(\sum_{\mu=1}^{M} \frac{X_{\mu i}}{\sqrt{N}} g_{\text{out},\mu}^t - \frac{X_{\mu i}^2}{N} \partial_\omega g_{\text{out},\mu}^t \hat{w}_i^t \right)$$

$$\Sigma_i^t = -\left(\sum_{\mu=1}^{M} \frac{X_{\mu i}^2}{N} \partial_\omega g_{\text{out},\mu}^t \right)^{-1}$$

Update of the estimated marginals $\hat{w}_i \in \mathbb{R}$ and $\hat{c}_i \in \mathbb{R}^+$:

$$\hat{w}_i^{t+1} = f_1^w(\Sigma_i^t, T_i^t)$$

$$\hat{c}_i^{t+1} = f_2^w(\Sigma_i^t, T_i^t)$$

$t = t + 1$

until Convergence on $\hat{w}$, $\hat{c}$.
Output: $\hat{w}$ and $\hat{c}$.

D. Step4: State evolution equations - AMP

We define the overlap parameters at time t, m^t, q^t, σ^t and Q. Especially m^t measures respectively correlation between the estimate and the planted solution.

$$\begin{cases} m^t \equiv \frac{1}{N}\sum_{i=1}^{N} \hat{w}_i^t (w_i^0)^\intercal \\ q^t \equiv \frac{1}{N}\sum_{i=1}^{N} \hat{w}_i^t (\hat{w}_i^t)^\intercal \end{cases} \quad \text{and} \quad \begin{cases} Q^0 \equiv \frac{1}{N}\sum_{i=1}^{N} w_i^0 (w_i^0)^\intercal \\ \sigma^t \equiv \frac{1}{N}\sum_{i=1}^{N} \hat{c}_i^t \end{cases} \tag{8}$$

The aim is to derive the asymptotic state evolution equations of these overlap parameters, starting with the relaxed BP equations Eq. II B. The idea is to compute the distributions of each variable to finally get the distribution of these overlaps.

The computation is shown in (Eq.IV C).

Bayes optimal case

In the Bayes optimal case, the student knows everything about the teacher, and in particular using Nishimori identities:

$$\begin{cases} P_{w^0} = P_w \\ P_{out}^0 = P_{out} \\ m^t = q^t \end{cases}$$

In this case, the state evolution equation read as scalar iterative equations:

$$\begin{cases} q^{t+1} = \mathbb{E}_\xi \left[f_0^w \left(\lambda^t[\xi], \sigma^t\right) f_1^w \left(\lambda^t[\xi], \sigma^t\right) f_1^w \left(\lambda^t[\xi], \sigma^t\right)^\intercal \right] \\ \hat{q}^t = \alpha \mathbb{E}_{y,\xi} \left[f_{out}(y, \omega^t[\xi], V^t) g_{out}(y, \omega^t[\xi], V^t) g_{out}(y, \omega^t[\xi], V^t) \right] \end{cases} \quad \text{and} \quad \begin{cases} \lambda^t[\xi] \equiv (\hat{q}^t)^{-1/2}\xi \\ \sigma^t \equiv (\hat{q}^t)^{-1} \\ \omega^t[\xi] = (q^t)^{1/2}\xi \\ V^t = Q^0 - q^t \end{cases} \tag{9}$$

with f_{out}, g_{out}, f_0^w and f_1^w defined in II E.

E. Update functions

Where the update functions read:

- f_w

$$\begin{cases} \tilde{P}_w(w,\lambda,\sigma) \equiv \dfrac{1}{f_0^w(\lambda,\sigma)} P_w(w) e^{-\frac{1}{2}w\sigma^{-1}w+\lambda\sigma^{-1}w} \\ f_0^w(\lambda,\sigma) = \displaystyle\int_{\mathbb{R}^K} dw P_w(w) e^{-\frac{1}{2}w\sigma^{-1}w+\lambda\sigma^{-1}w} \end{cases} \qquad \begin{cases} f_1^w(\lambda,\sigma) = \mathbb{E}_{\tilde{P}_w}\left[w\right] \\ f_2^w(\lambda,\sigma) = \mathbb{E}_{\tilde{P}_w}\left[ww\right] - f_1^w(f_1^w) \end{cases} \tag{10}$$

- f_{out}

$$\begin{cases} \tilde{P}_{out}(z;y,\omega,V) = \frac{1}{f_{out}(y,\omega,V)} P_{out}\left(y|\varphi_{out}\left(z;\Delta\right)\right) e^{-\frac{1}{2}(z-\omega)V^{-1}(z-\omega)} \\ f_{out}(y,\omega,V) = \mathcal{N}(V) \displaystyle\int_{\mathbb{R}} dz P_{out}\left(y|\varphi_{out}\left(z;\Delta\right)\right) e^{-\frac{1}{2}(z-\omega)V^{-1}(z-\omega)} \\ g_{out}(y,\omega,V) = \dfrac{1}{f_{out}} \dfrac{\partial f_{out}}{\partial \omega} = V^{-1}\mathbb{E}_{\tilde{P}_{out}}\left[z-\omega\right] \\ \partial_\omega g_{out}(y,\omega,V) = \dfrac{\partial g_{out}}{\partial \omega} = V^{-1}\mathbb{E}_{\tilde{P}_{out}}\left[(z-\omega)(z-\omega)\right] V^{-1} - V^{-1} - g_{out}^2(y,\omega,V) \end{cases} \tag{11}$$

III. Replicas computation

An other approach to tackle this high dimensional bayesian inference is to compute the averaged free entropy Φ. This is the central object of study in disordered systems and the entropy can be computed using the replica method.

We will show that in fact state evolution of the AMP algorithm can be obtained from the replica free entropy directly.

A. Step 1: Partition function and replica trick

1. Partition function, free entropy and average scenario

Let's define the partition function $\mathcal{Z}$ as the normalizing constant in (Eq.4), that reads integrating over $\underline{w}$:

$$\mathcal{Z}\left(\underline{y};\underline{\underline{X}}\right) \equiv \mathcal{P}\left(\underline{y};\underline{\underline{X}}\right) = \int_{\mathbb{R}^N} d\underline{w} P_w(\underline{w}) \int_{\mathbb{R}^M} d\underline{z} P_{out}\left(\underline{y}|\varphi_{out}\left(\underline{z},\Delta\right)\right) \delta\left(\underline{z} - \frac{1}{\sqrt{N}}\underline{w}\underline{\underline{X}}\right)$$

As we consider the average scenario, we need to average over the training set $\{\underline{y}, \underline{\underline{X}}\}$ and the planted solution $\underline{w}^0$. The averaged free entropy reads then:

$$\begin{aligned}
\Phi &= \lim_{N\to\infty} \frac{1}{N} \mathbb{E}_{\underline{y},\underline{w}^0,\underline{\underline{X}}}\left[\log \mathcal{Z}(\underline{y};\underline{\underline{X}})\right] \\
&= \lim_{N\to\infty} \mathbb{E}_{\underline{\underline{X}}}\left[\int_{\mathbb{R}^M} d\underline{y} \int_{\mathbb{R}^N} d\underline{w}^0 P_{w^0}(\underline{w}^0) \int_{\mathbb{R}^M} d\underline{z}^0 P^0_{out}\left(\underline{y}|\varphi^0_{out}\left(\underline{z}^0;\Delta\right)\right) \delta\left(\underline{z}^0 - \frac{1}{\sqrt{N}}\underline{w}^0\underline{\underline{X}}\right) \times ...\right. \\
&\quad \left. ... \times \log\left(\int_{\mathbb{R}^N} d\underline{w} P_w(\underline{w}) \int_{\mathbb{R}^M} d\underline{z} P_{out}\left(\underline{y}|\varphi_{out}\left(\underline{z},;\Delta\right)\right) \delta\left(\underline{z} - \frac{1}{\sqrt{N}}\underline{w}\underline{\underline{X}}\right)\right)\right]
\end{aligned}$$

2. Replica trick

This average of the logarithm is intractable. Instead we use the so-called *Replica trick*: $\log(z) = \lim_{n\to 0} \frac{z^n - 1}{n}$. Applying to the free entropy:

$$\begin{aligned}
\Phi &= \lim_{N\to\infty} \mathbb{E}_{\underline{\underline{X}}}\left[\phi\left(\underline{\underline{X}}\right)\right] = \lim_{N\to\infty} \frac{1}{N} \mathbb{E}_{\underline{y},\underline{w}^0,\underline{\underline{X}}}\left[\log \mathcal{Z}(\underline{y};\underline{\underline{X}})\right] \\
&= \lim_{n\to 0} \frac{1}{Nn} \log \mathbb{E}_{\underline{y},\underline{w}^0,\underline{\underline{X}}}\left[\mathcal{Z}(\underline{y};\underline{\underline{X}})^n\right]
\end{aligned}$$

B. Step 2: Average over disorder and interacting copies

a. Average of the moments

Remains to compute the averaged n-th moment of the partition function: $\mathbb{E}_{\underline{y},\underline{w}^0,\underline{\underline{X}}}\left[\mathcal{Z}(\underline{y};\underline{\underline{X}})^n\right]$. Note that computing $\mathcal{Z}^n$ corresponds to compute the partition function of $n \in \mathbb{N}$ non-interacting copies of the same system.

$$\mathbb{E}_{\underline{y},\underline{w}^0,\underline{\underline{X}}}\left[\mathcal{Z}(\underline{y};\underline{\underline{X}})^n\right] = \mathbb{E}_{\underline{\underline{X}}}\left[\int_{\mathbb{R}^M} d\underline{y} \int_{\mathbb{R}^N} d\underline{w}^0 P_{w^0}(\underline{w}^0) \int_{\mathbb{R}^M} d\underline{z}^0 P^0_{out}\left(\underline{y}|\varphi^0_{out}\left(\underline{z}^0;\Delta^0\right)\right)\delta(\underline{z}^0 - \underline{\underline{X}}\underline{w}^0)\right. \tag{12}$$

$$\left.\prod_{a=1}^n \int_{\mathbb{R}^N} d\underline{w}^a P_{w^a}(\underline{w}^a) \int_{\mathbb{R}^M} d\underline{z}^a P_{out}\left(\underline{y}|\varphi_{out}\left(\underline{z}^a;\Delta\right)\right)\delta(\underline{z}^a - \underline{\underline{X}}\underline{w}^a)\right] \tag{13}$$

b. Average over X

The first step is to take advantage of the *iid* property of the disorder distribution $P_x \sim \mathcal{N}(0,1)$.

Let's consider the variable $z^a_\mu = \displaystyle\sum_{i=1}^N X_{\mu i} w^a_i$. Using CLT, the vector z^a_μ follows a multivariate gaussian distribution such that:

$$\mathbb{E}_{\underline{\underline{X}}}\left[z^a_\mu\right] = 0 \quad \text{and} \quad \mathbb{E}_{\underline{\underline{X}}}\left[z^a_\mu z^b_\nu\right] = \frac{1}{N}\sum_{i=1}^N w^a_i w^b_i \delta_{\mu\nu} \tag{14}$$

COMMENT:

1. Replicas were identical and non-interacting as it was the copy of n identical systems. Strikingly, averaging over the disorder makes replicas interact
2. Besides it naturally introduces the order parameter, called the *overlap* $q_{ab} = \frac{1}{N}(\underline{w}^a)^\intercal \underline{w}^b$, that measures the correlation between replicas.

From the above average, $\underline{\mathbf{z}}$ follows a multivariate distribution of size $M \times (n+1)$: $P_z(\underline{\mathbf{z}}) = \frac{e^{-\frac{1}{2}\underline{\mathbf{z}}^\intercal \underline{\underline{\mathbf{\Sigma}}}^{-1}\underline{\mathbf{z}}}}{(2\pi)^{M(n+1)/2}\det(\underline{\underline{\mathbf{\Sigma}}})^{M/2}}$, with the matrix of overlaps as covariance matrix:

$$\left(\underline{\underline{\mathbf{\Sigma}}}\right)_{ab} = Q_{ab} \equiv \frac{1}{N}\sum_{i=1}^N w^a_i w^b_i = \frac{1}{N}(\underline{w}^a)^\intercal \underline{w}^b \tag{15}$$

The overlap being the natural order parameter, we introduce the change of variable multiplying by $1 = \int \prod_{(a,b)} dQ_{ab}\delta\left(NQ_{ab} - \sum_{i=1}^N w^a_i w^b_i\right)$. Hence denoting everything related to the teacher with the index $a=0$:

$$\mathbb{E}_{\underline{y},\underline{w}^0,\underline{\underline{X}}}\left[\mathcal{Z}(\underline{y};\underline{\underline{X}})^n\right] = \int \prod_{(a,b)} dQ_{ab} \tag{16}$$

$$\int d\underline{y} \prod_{a=0}^n \int d\underline{z}^a P^a_{out}\left(\underline{y}|\varphi^a_{out}\left(\underline{z}^a;\Delta^a\right)\right) \exp\left(-\frac{1}{2}\sum_{\mu=1}^M \sum_{ab} z^a_\mu z^b_\mu (\Sigma_{ab})^{-1} - \frac{M}{2}\log(\det\underline{\underline{\mathbf{\Sigma}}}) - \frac{M(n+1)}{2}\log(2\pi)\right) \tag{17}$$

$$\prod_{a=0}^n \int d\underline{w}^a P_{w^a}(\underline{w}^a) \prod_{(a,b)} \delta\left(NQ_{ab} - \sum_{i=1}^N w^a_i w^b_i\right) \tag{18}$$

Next we introduce the Fourier representation of the Dirac function: $\delta\left(NQ_{ab} - \sum_{i=1}^N w^a_i w^b_i\right) = \int d\hat{Q}_{ab} e^{\hat{Q}_{ab}\left(NQ_{ab} - \sum_{i=1}^N w^a_i w^b_i\right)}$. As P_w and P_{out} decouple in one direction, it allows to decouple the replicated partition functions as *prior* and *channel* terms:

$$\mathbb{E}_{\underline{y},\underline{w}^0,\underline{\underline{X}}}\left[\mathcal{Z}(\underline{y};\underline{\underline{X}})^n\right] = \int \prod_{(a,b)} dQ_{ab} \int \prod_{(a,b)} d\hat{Q}_{ab} \prod_{(a,b)} \exp\left(N\hat{Q}_{ab}Q_{ab}\right) \tag{19}$$

$$\int d\underline{y} \prod_{a=0}^{n} \int_{\mathbb{R}^M} d\underline{z}^a P_{out}^a\left(\underline{y}|\varphi_{out}^a\left(\underline{z}^a;\Delta^a\right)\right) \exp\left(-\frac{1}{2}\sum_{\mu=1}^{M}\sum_{ab} z_\mu^a z_\mu^b (\Sigma_{ab})^{-1} - \frac{M}{2}\log(\det\underline{\underline{\mathbf{\Sigma}}}) - \frac{M(n+1)}{2}\log(2\pi)\right) \tag{20}$$

$$\prod_{a=0}^{n} \int_{\mathbb{R}^N} d\underline{w}^a P_{w^a}(\underline{w}^a) \prod_{(a,b)} \exp\left(-\hat{Q}_{ab}\sum_{i=1}^{N} w_i^a w_i^b\right) \tag{21}$$

$$= \int \prod_{(a,b)} dQ_{ab} \int \prod_{(a,b)} d\hat{Q}_{ab} \prod_{(a,b)} \exp\left(N\hat{Q}_{ab}Q_{ab}\right) \tag{22}$$

$$\left[\int dy \int_{\mathbb{R}^{n+1}} \prod_{a=0}^{n} dz^a P_{out}^a\left(y|\varphi_{out}^a\left(z^a;\Delta^a\right)\right) \exp\left(-\frac{1}{2}\sum_{ab} z^a z^b (\Sigma_{ab})^{-1} - \frac{1}{2}\log(\det\underline{\underline{\mathbf{\Sigma}}}) - \frac{(n+1)}{2}\log(2\pi)\right)\right]^M \tag{23}$$

$$\left[\int_{\mathbb{R}^{n+1}} \prod_{a=0}^{n} dw^a P_{w^a}(w^a) \prod_{(a,b)} \exp\left(-\hat{Q}_{ab} w^a w^b\right)\right]^N \tag{24}$$

c. Conclusion

To conclude we need to use "non rigorous" tricks:

1. use an analytic continuation for $n \in \mathbb{N}$ that becomes $n \in \mathbb{R}$, before taking the limit $n \to 0$
2. exchange the limits $N \to \infty$ and $n \to 0$...
3. use a Laplace method of the integrals in the $N \to \infty$ limit

Doing so, the free entropy reads as a saddle point:

$$\begin{aligned}\Phi =& \frac{1}{N}\mathbb{E}_{\underline{y},\underline{w}^0,\underline{\underline{X}}}\left[\log\mathcal{Z}(\underline{y};\underline{\underline{X}})\right] = \lim_{n\to 0}\frac{1}{Nn}\log\mathbb{E}_{\underline{y},\underline{w}^0,\underline{\underline{X}}}\left[\mathcal{Z}(\underline{y},\underline{\underline{X}})^n\right] & (25)\\ &= \lim_{n\to 0}\frac{1}{n}\mathbf{extr}_{\underline{\underline{\mathbf{Q}}},\underline{\underline{\hat{\mathbf{Q}}}}}\left\{\mathcal{H}(\underline{\underline{\mathbf{Q}}},\underline{\underline{\hat{\mathbf{Q}}}})\right\} & (26)\end{aligned}$$

where:

$$\begin{cases}\mathcal{H}(\underline{\underline{\mathbf{Q}}},\underline{\underline{\hat{\mathbf{Q}}}}) & \equiv \sum_{ab}\hat{Q}_{ab}Q_{ab} + \log\left(\hat{\mathcal{I}}_w(\underline{\underline{\hat{\mathbf{Q}}}})\right) + \alpha\log\left(\hat{\mathcal{I}}_z(\underline{\underline{\mathbf{Q}}};\Delta\right)\\ \hat{\mathcal{I}}_w(\underline{\underline{\hat{\mathbf{Q}}}}) & \equiv \int_{\mathbb{R}^{n+1}}\prod_{a=0}^{n} dw^a P_{w^a}(w^a)\prod_{(a,b)}\exp\left(-\hat{Q}_{ab}w^a w^b\right)\\ \hat{\mathcal{I}}_z(\underline{\underline{\mathbf{Q}}};\Delta) & \equiv \int dy\int_{\mathbb{R}^{n+1}}\prod_{a=0}^{n} dz^a P_{out}^a\left(y|\varphi_{out}^a\left(z^a;\Delta\right)\right)\times\exp\left(-\frac{1}{2}\sum_{ab} z^a z^b(\Sigma^{ab})^{-1} - \frac{1}{2}\log(\det\underline{\underline{\mathbf{\Sigma}}}) - \frac{(n+1)}{2}\log(2\pi)\right)\end{cases} \tag{27}$$

Computing the free entropy is reduced to an optimization problem over two matrices $\underline{\underline{\mathbf{Q}}},\underline{\underline{\hat{\mathbf{Q}}}} \in \mathbb{R}^{(n+1)\times(n+1)}$.

C. Step 2: RS assumption

To take the limit $n \to 0$, we need an analytical expression for the potential $\mathcal{H}(\underline{\underline{\mathbf{Q}}}, \underline{\underline{\hat{\mathbf{Q}}}})$ as a function of n. Hence we assume a simple ansatz saying that the critical point of $\mathcal{H}$ is reached for a value $\underline{\underline{\mathbf{Q}}}^w$ and $\underline{\underline{\hat{\mathbf{Q}}}}^w$ which verify the Replica Symmetric assumption. More precisely, matrices $\underline{\underline{\mathbf{Q}}}, \underline{\underline{\hat{\mathbf{Q}}}}$ verify:

$$\underline{\underline{\mathbf{Q}}}^w = \begin{bmatrix} Q^0 & m & m & m \\ m & Q & q & q \\ m & q & Q & q \\ m & q & q & Q \end{bmatrix} \qquad \underline{\underline{\hat{\mathbf{Q}}}}^w = \begin{bmatrix} \underline{\underline{\hat{Q}}}^0 & \hat{m} & \hat{m} & \hat{m} \\ \hat{m} & \hat{Q} & \hat{q} & \hat{q} \\ \hat{m} & \hat{q} & \hat{Q} & \hat{q} \\ \hat{m} & \hat{q} & \hat{q} & \hat{Q} \end{bmatrix} \tag{28}$$

Note that we recover naturally the overlaps defined in AMP (Eq.8). These quantities have deep interpretation: m measures the correlation between the estimator and the planted solution, q is the correlation between all replicas, and Q, Q^0 are the norms of replicas and planted solutions.

$$\begin{cases} m = \frac{1}{N}(\underline{w}^0)^\intercal \underline{w}^a \text{ for } a \neq 0 \\ q = \frac{1}{N}(\underline{w}^a)^\intercal \underline{w}^b \text{ for } a \neq b \end{cases} \quad \text{and} \quad \begin{cases} Q^0 = \frac{1}{N}(\underline{w}^0)^\intercal \underline{w}^0 \\ Q = \frac{1}{N}(\underline{w}^a)^\intercal \underline{w}^a \text{ for } a \neq 0 \end{cases} \tag{29}$$

COMMENT: Note that in general this assumption is wrong. To check this, we need to study its stability (computing the eigenvalues of the hessian). It leads to Replica Symmetry Breaking (RSB)

Using this simple RS ansatz, we can now explicit the replica potential:

$$\mathcal{H}(\underline{\underline{\mathbf{Q}}}, \underline{\underline{\hat{\mathbf{Q}}}}) \equiv \sum_{ab} \hat{Q}_{ab} Q_{ab} + \log\left(\hat{\mathcal{I}}_w(\underline{\underline{\hat{\mathbf{Q}}}})\right) + \alpha \log\left(\hat{\mathcal{I}}_z(\underline{\underline{\mathbf{Q}}}; \Delta)\right) \tag{30}$$

We show details of the computation in (Eq.IV E). In the end we obtain an expression that is explicit in n:

$$\begin{cases} \hat{\mathcal{I}}_w(\underline{\underline{\hat{\mathbf{Q}}}}) = \int_{\mathbb{R}} D\xi \int_{\mathbb{R}} dw^0 P_{w^0}(w^0) e^{-(w^0)^\intercal \hat{Q}^0_w w^0} \left[\int_{\mathbb{R}} dw P_w(w) \exp\left(-\left[2w^0 \hat{m} w + w\left(\hat{Q} + \hat{q}\right) w - \xi \hat{q}^{1/2} w\right]\right)\right]^n \\ \hat{\mathcal{I}}_z(\underline{\underline{\mathbf{Q}}}; \Delta) = \int D\xi \int dy e^{-\frac{1}{2}\log(\det \underline{\underline{\boldsymbol{\Sigma}}}) - \frac{(n+1)}{2}\log(2\pi)} \int dz^0 P^0_{out}\left(y|\varphi^0_{out}\left(z^0; \Delta^0\right)\right) e^{-\frac{1}{2}\left(z^0 \Sigma^{-1}_{00} z^0\right)} \\ \left[\int dz P_{out}\left(y|\varphi_{out}\left(z; \Delta\right)\right) e^{-z^0 \Sigma^{-1}_{01} z - \frac{1}{2} z\left(\Sigma^{-1}_{11} - \Sigma^{-1}_{12}\right) z - \xi \Sigma^{-1/2}_{12} z}\right]^n \end{cases} \tag{31}$$

a. Check $n \to 0$ limit

We need to check that $\lim_{n\to 0} \mathbf{extr}_{\underline{\underline{\mathbf{Q}}}, \underline{\underline{\hat{\mathbf{Q}}}}} \left\{\mathcal{H}(\underline{\underline{\mathbf{Q}}}, \underline{\underline{\hat{\mathbf{Q}}}})\right\} \longrightarrow 0$. Doing so leads to ((Eq.IV F)):

$$\begin{cases} \hat{Q}^0 = 0 \\ Q^0 = \mathbb{E}_{P_{w^0}}\left[w^0 w^0\right] \end{cases}$$

D. Step 3: $n \to 0$ limit

The last step consists in dividing by n and taking the $n \to 0$ limit.

COMMENT:

- Note that we exchanged the order of the limit $N \to \infty$ and $n \to 0$: we use the Laplace method before taking $n \to 0$
- The computation is long, painful and does not give any insights. We skip the details and we refer to (Eq.IV G)
- However, in a few words the idea is to decouple interacting replicas using a Hubbard-Stratanovich / Gaussian transformation : $D\xi = \frac{e^{-\xi^2/2}}{\sqrt{2\pi}}$

$$\Phi = \lim_{n \to 0} \frac{1}{n} \mathbf{extr}_{\underline{\underline{\mathbf{Q}}}, \underline{\underline{\hat{\mathbf{Q}}}}} \left\{ \mathcal{H}(\underline{\underline{\mathbf{Q}}}, \underline{\underline{\hat{\mathbf{Q}}}}) \right\} \tag{32}$$

E. Conclusion - Bayes optimal setting

In the Bayes optimal setting, using Nishimori Identities,

$$\begin{cases} Q = Q^0 = \mathbb{E}_{P_w^0}\left[w^0 w^0\right] \\ \hat{Q} = \hat{Q}^0 = 0 \end{cases} \quad \text{and} \quad \begin{cases} q = m \\ \hat{q} = \hat{m} \end{cases} \quad \text{and} \quad \begin{cases} P_{w^0} = P_w \\ P_{out}^0 = P_{out} \\ , \end{cases}$$

the RS free entropy simplifies and reads as an optimization problem over two scalar parameters q and $\hat{q}$:

$$\Phi = \frac{1}{N}\mathbb{E}_{\underline{y},\underline{w}^0,\underline{\underline{X}}}\left[\log \mathcal{Z}(\underline{y};\underline{\underline{X}})\right] = \mathbf{extr}_{q,\hat{q}}\left\{-\frac{1}{2}q\hat{q} + \mathcal{I}_w\left(\hat{q}\right) + \alpha\mathcal{I}_z\left(Q^0, q; \Delta\right)\right\} \tag{33}$$

where

$$\begin{cases} \mathcal{I}_z\left(Q^0, q; \Delta\right) = \mathbb{E}_{y,\xi}\left[f_{out}\left(y, \omega[\xi], V; \Delta\right) \log f_{out}\left(y, \omega[\xi], V; \Delta\right)\right] \\ \mathcal{I}_w\left(\hat{q}\right) = \mathbb{E}_{\xi}\left[f_0^w\left(\lambda[\xi], \sigma\right) \log f_0^w\left(\lambda[\xi], \sigma\right)\right] \end{cases} \quad \text{and} \quad \begin{cases} \sigma^{-1}\lambda[\xi] = \hat{q}^{1/2}\xi \\ \sigma = \left(\hat{q}\right)^{-1} \\ \omega[\xi] = q^{1/2}\xi \\ V = \left(Q^0 - q\right) \end{cases} \tag{34}$$

with f_{out} and f_0^w defined in (Eq.II E)

a. About proof and theorem

This free entropy can be proved and legitimate all the non rigorous tricks in the replica method. Cf *Bayati-Montanari 2001*, *Rangan 2010*

F. State evolution and consistence with replicas

Taking the saddle point of (Eq.40), and using integration by part, we obtain:

$$\begin{cases} q = \mathbb{E}_{\xi}\left[f_0^w\left(\lambda[\xi], \sigma\right) f_1^w\left(\lambda[\xi], \sigma\right) f_1^w\left(\lambda[\xi], \sigma\right)\right] \\ \hat{q} = \alpha\mathbb{E}_{y,\xi}\left[f_{out}(y, \omega[\xi], V) g_{out}(y, \omega[\xi], V) g_{out}(y, \omega[\xi], V)\right] \end{cases}$$

This formulation is equivalent to the one we obtain with AMP(Eq.9), except we do not have the time indices... ! Hence the state evolution of the AMP algorithm does a gradient descent of the RS free entropy (in the Bayes optimal case).

IV. Appendices

A. Towards relaxed BP

The term inside P_{out} can be decouple using its Fourrier transform:

$$P_{\text{out}}\left(y_\mu | \frac{1}{\sqrt{N}}\sum_{j=1}^{N} X_{\mu j} w_j\right) = \frac{1}{(2\pi)^{K/2}} \int_{\mathbb{R}} d\xi \exp\left(i\xi^\intercal \left(\frac{1}{\sqrt{N}}\sum_{j=1}^{N} X_{\mu j} w_j\right) \hat{P}_{\text{out}}(y_\mu, \xi)\right).$$

Injecting this representation in the BP equations, (5) becomes:

$$\tilde{m}^t_{\mu\to i}(w_i) = \frac{1}{(2\pi)^{K/2}\mathcal{Z}_{\mu\to i}} \int_{\mathbb{R}} d\xi \hat{P}_{\text{out}}(y_\mu, \xi) \exp\left(i\xi^\intercal \frac{1}{\sqrt{N}} X_{\mu i} w_i\right) \times \prod_{j\neq i}^{N} \underbrace{\int_{\mathbb{R}} dw_j m^t_{j\to\mu}(w_j) \exp\left(i\xi^\intercal \frac{1}{\sqrt{N}} X_{\mu j} w_j\right)}_{\equiv I_j}$$

In the limit $n\to\infty$ the term I_j can be easily expanded and expressed using $\hat{w}$ and $\hat{c}$:

$$I_j = \int_{\mathbb{R}} dw_j m^t_{j\to\mu}(w_j) \exp\left(i\xi^\intercal \frac{X_{\mu j}}{\sqrt{N}} w_j\right) \simeq \exp\left(i\frac{X_{\mu j}}{\sqrt{N}}\xi^\intercal \hat{w}^t_{j\to\mu} - \frac{1}{2}\frac{X^2_{\mu j}}{N}\xi^\intercal \hat{c}^t_{j\to\mu}\xi\right).$$

And finally using the inverse Fourier transform:

$$\begin{aligned}
\tilde{m}^t_{\mu\to i}(w_i) &= \frac{1}{(2\pi)\mathcal{Z}_{\mu\to i}} \int_{\mathbb{R}} dz P_{\text{out}}(y_\mu, z) \int_{\mathbb{R}} d\xi e^{-i\xi^\intercal z} e^{iX_{\mu i}\xi^\intercal w_i} \\
&\quad \times \prod_{j\neq i}^{N} \exp\left(i\frac{X_{\mu j}}{\sqrt{N}}\xi^\intercal \hat{w}^t_{j\to\mu} - \frac{1}{2}\frac{X^2_{\mu j}}{N}\xi^\intercal \hat{c}^t_{j\to\mu}\xi\right) \\
&= \frac{1}{(2\pi)\mathcal{Z}_{\mu\to i}} \int_{\mathbb{R}} dz P_{\text{out}}(y_\mu, z) \int_{\mathbb{R}} d\xi e^{-i\xi^\intercal z} e^{iX_{\mu i}\xi^\intercal w_i} e^{i\xi^\intercal \sum_{j\neq i}^{N} \frac{X_{\mu j}}{\sqrt{N}} \hat{w}^t_{j\to\mu}} e^{-\frac{1}{2}\xi^\intercal \sum_{j\neq i}^{N} \frac{X^2_{\mu j}}{N} \hat{c}^t_{j\to\mu}\xi} \\
&= \frac{1}{(2\pi)\mathcal{Z}_{\mu\to i}} \int_{\mathbb{R}} dz P_{\text{out}}(y_\mu, z) \sqrt{\frac{(2\pi)}{\det(V^t_{i\mu})}} \underbrace{e^{-\frac{1}{2}\left(z - \frac{X_{\mu i}}{\sqrt{N}} w_i - \omega^t_{i\mu}\right)^\intercal (V^t_{i\mu})^{-1} \left(z - \frac{X_{\mu i}}{\sqrt{N}} w_i - \omega^t_{i\mu}\right)}}_{\equiv H_{i\mu}}
\end{aligned}$$

where we defined the mean and variance, depending on the node i:

$$\begin{cases} \omega^t_{i\mu} \equiv \frac{1}{\sqrt{N}} \sum_{j\neq i}^{N} X_{\mu j} \hat{w}^t_{j\to\mu} \\ V^t_{i\mu} \equiv \frac{1}{N} \sum_{j\neq i}^{N} X^2_{\mu j} \hat{c}^t_{j\to\mu} \end{cases}$$

Again, in the limit $n\to\infty$, the term $H_{i\mu}$ can be expanded:

$$\begin{aligned}
H_{i\mu} \simeq e^{-\frac{1}{2}\left(z-\omega^t_{i\mu}\right)^\intercal (V^t_{i\mu})^{-1}\left(z-\omega^t_{i\mu}\right)} &\left(1 + \frac{X_{\mu i}}{\sqrt{N}} w_i^\intercal (V^t_{i\mu})^{-1}(z-\omega^t_{i\mu}) - \frac{1}{2}\frac{X^2_{\mu i}}{N} w_i^\intercal (V^t_{i\mu})^{-1} w_i \right. \\
&\left. + \frac{1}{2}\frac{X^2_{\mu i}}{N} w_i^\intercal (V^t_{i\mu})^{-1}(z-\omega^t_{i\mu})(z-\omega^t_{i\mu})^\intercal (V^t_{i\mu})^{-1} w_i\right).
\end{aligned}$$

Gathering all pieces, the message $\tilde{m}_{\mu\to i}$ can be expressed using definitions of g_{out} and $\partial_\omega g_{\text{out}}$:

$$
\begin{aligned}
\tilde{m}^t_{\mu\to i}(w_i) \sim \frac{1}{\mathcal{Z}_{\mu\to i}} &\left\{1 + \frac{X_{\mu i}}{\sqrt{N}} w_i^\intercal g_{\text{out}}(\omega^t_{i\mu}, y_\mu, V^t_{i\mu}) + \frac{1}{2}\frac{X^2_{\mu i}}{N} w_i^\intercal g_{\text{out}} g_{\text{out}}^\intercal(\omega^t_{i\mu}, y_\mu, V^t_{i\mu}) w_i + \right. \\
&\left. \frac{1}{2}\frac{X^2_{\mu i}}{N} w_i^\intercal \partial_\omega g_{\text{out}}(\omega^t_{i\mu}, y_\mu, V^t_{i\mu}) w_i \right\} \\
&= \frac{1}{\mathcal{Z}_{\mu\to i}} \left\{1 + w_i^\intercal B^t_{\mu\to i} + \frac{1}{2} w_i^\intercal B^t_{\mu\to i} (B^t_{\mu\to i})^\intercal (w_i) - \frac{1}{2} w_i^\intercal A^t_{\mu\to i} w_i \right\} \\
&= \sqrt{\frac{\det(A^t_{\mu\to i})}{(2\pi)}} \exp\left(-\frac{1}{2}\left(w_i^\intercal - (A^t_{\mu\to i})^{-1} B^t_{\mu\to i}\right)^\intercal A^t_{\mu\to i} \left(w_i^\intercal - (A^t_{\mu\to i})^{-1} B^t_{\mu\to i}\right)\right)
\end{aligned}
$$

with the following definitions of $A_{\mu\to i}$ and $B_{\mu\to i}$:

$$
\begin{cases}
B^t_{\mu\to i} \equiv \frac{X_{\mu i}}{\sqrt{N}} g_{\text{out}}(\omega^t_{i\mu}, y_\mu, V^t_{i\mu}) \\
A^t_{\mu\to i} \equiv -\frac{X^2_{\mu i}}{N} \partial_\omega g_{\text{out}}(\omega^t_{i\mu}, y_\mu, V^t_{i\mu})
\end{cases}
$$

Using the set of BP equations (5), we can close the set of equations only over $\{m_{i\to\mu}\}_{i\mu}$:

$$
m^{t+1}_{i\to\mu}(w_i) = \frac{1}{\mathcal{Z}_{i\to\mu}} P_0(w_i) \prod_{\nu\neq\mu}^{M} \sqrt{\frac{\det(A^t_{\nu\to i})}{(2\pi)}} e^{-\frac{1}{2}\left(w_i - (A^t_{\nu\to i})^{-1} B^t_{\nu\to i}\right)^\intercal A^t_{\nu\to i}\left(w_i - (A^t_{\nu\to i})^{-1} B^t_{\nu\to i}\right)}.
$$

In the end, computing the mean and variance of the product of gaussians, the messages are updated using f_w and f_c:

$$
\begin{cases}
\hat{w}^{t+1}_{i\to\mu} = f_w(\Sigma^t_{\mu\to i}, T^t_{\mu\to i}) \\
\hat{c}^{t+1}_{i\to\mu} = f_c(\Sigma^t_{\mu\to i}, T^t_{\mu\to i})
\end{cases}
\qquad
\begin{cases}
\Sigma^t_{\mu\to i} = \left(\sum_{\nu\neq\mu}^{M} A^t_{\nu\to i}\right)^{-1} \\
T^t_{\mu\to i} = \Sigma^t_{\mu\to i}\left(\sum_{\nu\neq\mu}^{M} B^t_{\nu\to i}\right)
\end{cases}
$$

B. Towards AMP

Let's now expand the previous messages Eq. II B, making appear these new target-independent messages:

a. $\Sigma^t_{\mu\to i}$

$$\Sigma^t_{\mu\to i} = \left(\sum_{\nu\neq\mu}^M A^t_{\nu\to i}\right)^{-1} = \left(\sum_{\nu=1}^M A^t_{\nu\to i} - A^t_{\mu\to i}\right)^{-1} = \left(\sum_{\nu=1}^M A^t_{\nu\to i}\left(I_{K\times K} - \left(\sum_{\nu=1}^M A^t_{\nu\to i}\right)^{-1} A^t_{\mu\to i}\right)\right)^{-1}$$
$$= \left(I_{K\times K} - \left(\sum_{\nu=1}^M A^t_{\nu\to i}\right)^{-1} A^t_{\mu\to i}\right)^{-1}\left(\sum_{\nu=1}^M A^t_{\nu\to i}\right)^{-1} = \underbrace{\left(I_{K\times K} - \Sigma^t_i A^t_{\mu\to i}\right)^{-1}}_{\simeq I_{K\times K}+\Sigma^t_i A^t_{\mu\to i}+\mathcal{O}(n^{-1})}\Sigma^t_i \simeq \Sigma^t_i + \mathcal{O}\left(\frac{1}{N}\right)$$

b. $T^t_{\mu\to i}$

$$T^t_{\mu\to i} = \Sigma^t_{\mu\to i}\left(\sum_{\nu\neq\mu}^M B^t_{\nu\to i}\right) = \left(\Sigma^t_i + \mathcal{O}\left(\frac{1}{N}\right)\right)\left(\sum_{\nu=1}^M B^t_{\nu\to i} - B^t_{\mu\to i}\right)$$
$$= T^t_i - \Sigma^t_i B^t_{\mu\to i} + \mathcal{O}\left(\frac{1}{N}\right)$$

c. $\hat{w}^{t+1}_{i\to\mu}$

$$\hat{w}^{t+1}_{i\to\mu} = f_w(\Sigma^t_{\mu\to i}, T^t_{\mu\to i}) = f_w\left(\Sigma^t_i, T^t_i - \Sigma^t_i B^t_{\mu\to i}\right) + \mathcal{O}\left(\frac{1}{N}\right)$$
$$\simeq f_w\left(\Sigma^t_i, T^t_i\right) - \left.\frac{df_w}{dT}\right|_{(\Sigma^t_i, T^t_i)} \Sigma^t_i B^t_{\mu\to i}$$
$$= \underbrace{f_w\left(\Sigma^t_i, T^t_i\right)}_{=\hat{w}^{t+1}_i} - \left(\Sigma^t_i\right)^{-1}\underbrace{f_c\left(\Sigma^t_i, T^t_i\right)\Sigma^t_i}_{=\hat{c}^{t+1}_i}\quad \underbrace{B^t_{\mu\to i}}_{\simeq \frac{X_{\mu i}}{\sqrt{N}} g_{\text{out}}(\omega^t_\mu, y_\mu, V^t_\mu)}$$
$$= \hat{w}^{t+1}_i - \frac{X_{\mu i}}{\sqrt{N}}\left(\Sigma^t_i\right)^{-1}\hat{c}^{t+1}_i \Sigma^t_i g_{\text{out}}(\omega^t_\mu, y_\mu, V^t_\mu) + \mathcal{O}\left(\frac{1}{N}\right)$$

d. $\hat{c}^{t+1}_{i\to\mu}$

Let's denote for convenience, $\mathcal{E} = \left(\Sigma^t_i\right)^{-1}\hat{c}^{t+1}_i \Sigma^t_i g_{\text{out}}(\omega^t_\mu, y_\mu, V^t_\mu)$. Then

$$\hat{c}^{t+1}_{i\to\mu} = \mathbb{E}_{\tilde{P}_0}\left[\hat{w}^t_{i\to\mu}(\hat{w}^t_{i\to\mu})^\intercal\right] - \mathbb{E}_{\tilde{P}_0}\left[\hat{w}^t_{i\to\mu}\right]\mathbb{E}_{\tilde{P}_0}\left[\hat{w}^t_{i\to\mu}\right]^\intercal$$
$$= \mathbb{E}_{\tilde{P}_0}\left[\left(\hat{w}^t_i - \frac{X_{\mu i}}{\sqrt{N}}\mathcal{E}\right)\left(\hat{w}^t_i - \frac{X_{\mu i}}{\sqrt{N}}\mathcal{E}\right)^\intercal\right] - \mathbb{E}_{\tilde{P}_0}\left[\hat{w}^t_i - \frac{X_{\mu i}}{\sqrt{N}}\mathcal{E}\right]\mathbb{E}_{\tilde{P}_0}\left[\hat{w}^t_i - \frac{X_{\mu i}}{\sqrt{N}}\mathcal{E}\right]^\intercal$$
$$= \mathbb{E}_{\tilde{P}_0}\left[\hat{w}^t_i(\hat{w}^t_i)^\intercal\right] - \mathbb{E}_{\tilde{P}_0}\left[\hat{w}^t_i\right]\mathbb{E}_{\tilde{P}_0}\left[\hat{w}^t_i\right]^\intercal + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right) = \hat{c}^{t+1}_i + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right)$$

e. $g_{\text{out}}(\omega^t_{i\mu}, y_\mu, V^t_{i\mu})$

$$
\begin{aligned}
g_{\text{out}}(\omega^t_{i\mu}, y_\mu, V^t_{i\mu}) &= g_{\text{out}}\left(\omega^t_\mu - \frac{X_{\mu i}}{\sqrt{N}}\hat{w}^t_{i\to\mu}, y_\mu, V^t_\mu - \frac{X^2_{\mu i}}{N}\hat{c}^t_{i\to l}\right) \\
&= g_{\text{out}}\left(\omega^t_\mu, y_\mu, V^t_\mu\right) - \frac{X_{\mu i}}{\sqrt{N}}\frac{\partial g_{\text{out}}}{\partial\omega}\left(\omega^t_\mu, y_\mu, V^t_\mu\right) \underbrace{\hat{w}^t_{i\to\mu}}_{=\hat{w}^t_i+\mathcal{O}\left(\frac{1}{\sqrt{N}}\right)} + \mathcal{O}\left(\frac{1}{N}\right) \\
&= g_{\text{out}}\left(\omega^t_\mu, y_\mu, V^t_\mu\right) - \frac{X_{\mu i}}{\sqrt{N}}\frac{\partial g_{\text{out}}}{\partial\omega}\left(\omega^t_\mu, y_\mu, V^t_\mu\right)\hat{w}^t_i + \mathcal{O}\left(\frac{1}{N}\right)
\end{aligned}
$$

f. V^t_μ

$$
V^t_\mu = \sum_{i=1}^N \frac{X^2_{\mu i}}{N}\hat{c}^t_{i\to l} = \sum_{i=1}^N \frac{X^2_{\mu i}}{N}\hat{c}^t_i + \mathcal{O}\left(\frac{1}{n^{3/2}}\right)
$$

g. ω^t_μ

$$
\begin{aligned}
\omega^t_\mu &= \sum_{i=1}^N \frac{X_{\mu i}}{\sqrt{N}}\hat{w}^t_{i\to\mu} = \sum_{i=1}^N \frac{X_{\mu i}}{\sqrt{N}}\left(\hat{w}^t_i - X_{\mu i}\left(\Sigma^{t-1}_i\right)^{-1}\hat{c}^t_i\Sigma^{t-1}_i g_{\text{out}}(\omega^{t-1}_\mu, y_\mu, V^{t-1}_\mu) + \mathcal{O}\left(\frac{1}{N}\right)\right) \\
&= \sum_{i=1}^N \frac{X_{\mu i}}{\sqrt{N}}\hat{w}^t_i - \sum_{i=1}^N \frac{X^2_{\mu i}}{N}\left(\Sigma^{t-1}_i\right)^{-1}\hat{c}^t_i\Sigma^{t-1}_i g_{\text{out}}(\omega^{t-1}_\mu, y_\mu, V^{t-1}_\mu) + \mathcal{O}\left(\frac{1}{n^{3/2}}\right)
\end{aligned}
$$

h. $\left(\Sigma^t_i\right)^{-1}$

$$
\left(\Sigma^t_i\right)^{-1} = \sum_{\mu=1}^M A^t_{\mu\to i} = -\sum_{\mu=1}^M X^2_{\mu i}\partial_\omega g_{\text{out}}(\omega^t_{i\mu}, y_\mu, V^t_{i\mu}) = -\sum_{\mu=1}^M X^2_{\mu i}\partial_\omega g_{\text{out}}(\omega^t_\mu, y_\mu, V^t_\mu) + \mathcal{O}\left(\frac{1}{n^{3/2}}\right)
$$

i. T^t_i

$$
\begin{aligned}
T^t_i &= \Sigma^t_i\left(\sum_{\mu=1}^M B^t_{\mu\to i}\right) = \Sigma^t_i\sum_{\mu=1}^M \frac{X_{\mu i}}{\sqrt{N}} g_{\text{out}}(\omega^t_{i\mu}, y_\mu, V^t_{i\mu}) \\
&= \Sigma^t_i\sum_{\mu=1}^M \frac{X_{\mu i}}{\sqrt{N}}\left(g_{\text{out}}\left(\omega^t_\mu, y_\mu, V^t_\mu\right) - \frac{X_{\mu i}}{\sqrt{N}}\frac{\partial g_{\text{out}}}{\partial\omega}\left(\omega^t_\mu, y_\mu, V^t_\mu\right)\hat{w}^t_i + \mathcal{O}\left(\frac{1}{N}\right)\right) \\
&= \Sigma^t_i\left(\sum_{\mu=1}^M \frac{X_{\mu i}}{\sqrt{N}} g_{\text{out}}\left(\omega^t_\mu, y_\mu, V^t_\mu\right) - \frac{X^2_{\mu i}}{N}\frac{\partial g_{\text{out}}}{\partial\omega}\left(\omega^t_\mu, y_\mu, V^t_\mu\right)\hat{w}^t_i\right) + \mathcal{O}\left(\frac{1}{n^{3/2}}\right)
\end{aligned}
$$

C. State evolution from AMP

1. Messages distribution

In order to get the state evolution of the overlap parameters, we need to compute the distribution of $\Sigma^t_{\mu\to i}$ and $T^t_{\mu\to i}$. Besides, we recall that in our model $y_\mu = \varphi^0_{out}\left(\frac{1}{\sqrt{N}} w^0 X_\mu, A\right)$. We define $z_\mu \equiv \frac{1}{\sqrt{N}} w^0 X_\mu = \frac{1}{\sqrt{N}} \sum_{i=1}^N X_{\mu i} w^0_i$ and $z_{\mu\to i} \equiv \frac{1}{\sqrt{N}} \sum_{j\neq i}^N X_{\mu j} w^0_j$. And it useful to recall $\mathbb{E}_X[X_{\mu i}] = 0$ and $\mathbb{E}_X[X^2_{\mu i}] = 1$.

- $\omega^t_{\mu\to i}$

Under Belief Propagation assumption messages are independent, $\omega^t_{\mu\to i}$ is thus the sum of independent variables and follows a gaussian distribution. Let's compute the first two moments, using expansions of the Approximate Message Passing equations:

$$\mathbb{E}_X\left[\omega^t_{\mu\to i}\right] = \frac{1}{\sqrt{N}} \sum_{j\neq i}^N \mathbb{E}_X\left[X_{\mu j}\right] \hat{w}^t_{j\to\mu} = 0$$

$$\begin{aligned}
\mathbb{E}_X\left[\omega^t_{\mu\to i}(\omega^t_{\mu\to i})^\intercal\right] &= \frac{1}{N} \sum_{j\neq i,k\neq i}^N \mathbb{E}_X\left[X_{\mu j}X_{\mu k}\right] \hat{w}^t_{j\to\mu}(\hat{w}^t_{k\to\mu})^\intercal = \sum_{j\neq i}^N \mathbb{E}_X\left[X^2_{\mu j}\right] \hat{w}_{j\to\mu}(\hat{w}_{j\to\mu})^\intercal \\
&= \frac{1}{N} \sum_{j\neq i}^N \hat{w}^t_{j\to\mu}(\hat{w}^t_{j\to\mu})^\intercal = \frac{1}{N} \sum_{i=1}^N \hat{w}^t_i(\hat{w}^t_i)^\intercal + \mathcal{O}\left(1/n^{3/2}\right) = q^t
\end{aligned}$$

- z_μ

$$\mathbb{E}_X\left[z_\mu\right] = \frac{1}{\sqrt{N}} \sum_{i=1}^N \mathbb{E}_X\left[X_{\mu i}\right] w^0_i = 0$$

$$\mathbb{E}_X\left[z_\mu z_\mu^\intercal\right] = \frac{1}{N} \sum_{j=1,k=1}^N \mathbb{E}_X\left[X_{\mu j}X_{\mu k}\right] w^0_j(W^0_k)^\intercal = \frac{1}{N} \sum_{i=1}^N w^0_i(w^0_i)^\intercal = Q$$

- z_μ *and* $\omega^t_{\mu\to i}$

$$\begin{aligned}
\mathbb{E}_X\left[\omega^t_{\mu\to i} z_\mu^\intercal\right] &= \frac{1}{N} \sum_{j\neq i,k=1}^N \mathbb{E}_X\left[X_{\mu j}X_{\mu k}\right] \hat{w}^t_{j\to\mu}(W^0_k)^\intercal = \frac{1}{N} \sum_{j\neq i}^N \hat{w}^t_{j\to\mu}(w^0_j)^\intercal \\
&= \frac{1}{N} \sum_{i=1}^N \hat{w}^t_i(w^0_i)^\intercal + \mathcal{O}\left(1/n^{3/2}\right) = m^t
\end{aligned}$$

Hence z_μ and $\omega^t_{\mu\to i}$ follow a Gaussian distribution with correlation matrix $\mathbf{Q}^t = \begin{bmatrix} Q & m^t \\ m^t & q^t \end{bmatrix}$.

- $V_{\mu\to i}$
concentrates around its mean:

$$\mathbb{E}_X\left[V^t_{\mu\to i}\right] = \frac{1}{N}\sum_{j\neq i}^{N}\mathbb{E}_X\left[X^2_{\mu j}\right]\hat{c}^t_{j\to\mu} = \frac{1}{N}\sum_{j\neq i}^{N}\hat{c}^t_{j\to\mu} = \frac{1}{N}\sum_{i}^{N}\hat{c}^t_i + \mathcal{O}\left(1/n^{3/2}\right) = \sigma^t$$

Let's define other order parameters, that will appear in the following:

$$\begin{cases}\hat{q}^t = \alpha\mathbb{E}_{\omega,z,A}\left[g_{out}(\omega,\varphi^0_{out}(z,A),\sigma^t)g_{out}(\omega,\varphi^0_{out}(z,A),\sigma^t)^\intercal\right]\\ \hat{m}^t = \alpha\mathbb{E}_{\omega,z,A}\left[\partial_z g_{out}(\omega,\varphi^0_{out}(z,A),\sigma^t)\right]\\ \hat{\chi}^t = \alpha\mathbb{E}_{\omega,z,A}\left[-\partial_\omega g_{out}(\omega,\varphi^0_{out}(z,A),\sigma^t)\right]\end{cases}$$

- $T^t_{\mu\to i}$
can be expanded around $z_{\mu\to i}$:

$$\begin{aligned}\left(\Sigma^t_{\mu\to i}\right)^{-1}T^t_{\mu\to i} &= \left(\sum_{\nu\neq\mu}^{M}B^t_{\nu\to i}\right) = \left(\sum_{\nu\neq\mu}^{M}\frac{1}{\sqrt{N}}X_{\nu i}g_{\text{out}}(\omega^t_{\nu\to i},\varphi^0_{out}\left(\frac{1}{\sqrt{N}}\sum_{j\neq i}^{N}X_{\mu j}w^0_j + X_{\mu i}w^0_i, A\right), V^t_{\nu\to i})\right)\\ &= \left(\sum_{\nu\neq\mu}^{M}\frac{1}{\sqrt{N}}X_{\nu i}g_{\text{out}}(\omega^t_{\nu\to i},\varphi^0_{out}\left(z_{\mu\to i},A\right),V^t_{\nu\to i})\right) + \left(\sum_{\nu\neq\mu}^{M}\frac{1}{N}X^2_{\nu i}\partial_z g_{\text{out}}(\omega^t_{\nu\to i},\varphi^0_{out}\left(z_{\mu\to i},A\right),V^t_{\nu\to i})\right)w^0_i\end{aligned}$$

- $\Sigma^t_{\mu\to i}$

$$\begin{aligned}\left(\Sigma^t_{\mu\to i}\right)^{-1} &= \sum_{\nu\neq\mu}^{M}A^t_{\nu\to i} = -\sum_{\nu\neq\mu}^{M}\frac{1}{N}X^2_{\nu i}\partial_\omega g_{\text{out}}(\omega^t_{\nu\to i},Y_\nu,V^t_{\nu\to i})\\ &= -\sum_{\nu\neq\mu}^{M}\frac{1}{N}X^2_{\nu i}\partial_\omega g_{\text{out}}(\omega^t_{\nu\to i},\varphi^0_{out}\left(z_{\nu\to i}\right),V^t_{\nu\to i}) + \mathcal{O}\left(1/n^{3/2}\right)\end{aligned}$$

Hence the first moments of the variables $\Sigma^t_{\mu\to i}$ and $T^t_{\mu\to i}$ read:

$$\begin{cases}\mathbb{E}_{\omega,z,A,X}\left[\left(\Sigma^t_{\mu\to i}\right)^{-1}T^t_{\mu\to i}\right] = \hat{m}^t w^0_i\\ \mathbb{E}_{\omega,z,A,X}\left[\left(\Sigma^t_{\mu\to i}\right)^{-1}T^t_{\mu\to i}\left(T^t_{\mu\to i}\right)^\intercal\left(\Sigma^t_{\mu\to i}\right)^{-1}\right] = \hat{q}^t\\ \mathbb{E}_{\omega,z,A,X}\left[\left(\Sigma^t_{\mu\to i}\right)^{-1}\right] = \hat{\chi}^t\end{cases}$$

And finally $T^t_{\mu\to i}\sim(\hat{\chi}^t)^{-1}\left(\hat{m}^t w^0_i + (\hat{q}^t)^{1/2}\xi\right)$ with $\xi\sim\mathcal{N}(0,\mathbb{1})$ and $\left(\Sigma^t_{\mu\to i}\right)^{-1}\sim(\hat{\chi}^t)^{-1}$

2. State evolution equations - Non Bayes optimal case

We define the following quantities:

$$\begin{aligned}T^t[w^0,\xi] &\equiv (\hat{\chi}^t)^{-1}\left(\hat{m}^t w^0 + (\hat{q}^t)^{1/2}\xi\right)\\ \Sigma^t &\equiv (\hat{\chi}^t)^{-1}\end{aligned}$$

Thus the state evolution equations read:

$$\begin{cases} m^{t+1} = \frac{1}{N}\sum_{i=1}^{N} \hat{w}_i^{t+1} (w_i^0)^\intercal \underset{N\to\infty}{\longrightarrow} \mathbb{E}_{w^0,\xi}\left[f_w\left(\Sigma^t, T^t[w^0,\xi]\right)\left(w^0\right)^\intercal \right] \\ q^{t+1} = \frac{1}{N}\sum_{i=1}^{N} \hat{w}_i^{t+1} (\hat{w}_i^{t+1})^\intercal \underset{N\to\infty}{\longrightarrow} \mathbb{E}_{w^0,\xi}\left[f_w\left(\Sigma^t, T^t[w^0,\xi]\right) f_w\left(\Sigma^t, T^t[w^0,\xi]\right)^\intercal \right] \\ \sigma^{t+1} = \frac{1}{N}\sum_{i=1}^{N} \hat{c}_i^{t+1} \underset{N\to\infty}{\longrightarrow} \mathbb{E}_{w^0,\xi}\left[f_c\left(\Sigma^t, T^t[w^0,\xi]\right)\right] \end{cases}$$

and

$$\begin{cases} \hat{q}^t \underset{N\to\infty}{\longrightarrow} \alpha\mathbb{E}_{\omega,z,A}\left[g_{out}(\omega, \varphi^0_{out}(z,A), \sigma^t) g_{out}(\omega, \varphi^0_{out}(z,A), \sigma^t)^\intercal \right] \\ = \alpha \int dP_A(A) \int dz d\omega \mathcal{N}\left(z,\omega; 0, \mathbf{Q}_w^t\right) g_{out}(\omega, \varphi^0_{out}(z,A), \sigma^t) g_{out}(\omega, \varphi^0_{out}(z,A), \sigma^t)^\intercal \\ \hat{m}^t \underset{N\to\infty}{\longrightarrow} \alpha\mathbb{E}_{\omega,z,A}\left[\partial_z g_{out}(\omega, \varphi^0_{out}(z,A), \sigma^t) \right] \\ = \alpha \int dP_A(A) \int dz d\omega \mathcal{N}\left(z,\omega; 0, \mathbf{Q}_w^t\right) \partial_z g_{out}(\omega, \varphi^0_{out}(z,A), \sigma^t) \\ \hat{\chi}^t \underset{N\to\infty}{\longrightarrow} \alpha\mathbb{E}_{\omega,z,A}\left[-\partial_\omega g_{out}(\omega, \varphi^0_{out}(z,A), \sigma^t) \right] \\ = -\alpha \int dP_A(A) \int dz d\omega \mathcal{N}\left(z,\omega; 0, \mathbf{Q}_w^t\right) \partial_\omega g_{out}(\omega, \varphi^0_{out}(z,A), \sigma^t) \end{cases}$$

3. State evolution equations - Bayes optimal case

In the bayes optimal case, $P_{w^0} = P_W$, $\varphi^0_{out} = \varphi_{out}$, $m^t = q^t$ and $\hat{q}^t = \hat{m}^t = \hat{\chi}^t$, $\sigma^t = Q - q^t$ and

$$\begin{aligned} T^t[w^0,\xi] &\equiv w^0 + (\hat{q}^t)^{-1/2}\xi \\ \Sigma^t &\equiv (\hat{q}^t)^{-1} . \end{aligned}$$

Thus the state evolution equations simplify:

$$\begin{cases} q^{t+1} \underset{N\to\infty}{\longrightarrow} \alpha\mathbb{E}_{w^0,\xi}\left[f_w\left(\Sigma^t, T^t[w^0,\xi]\right) f_w\left(\Sigma^t, T^t[w^0,\xi]\right)^\intercal \right] \\ \hat{q}^t \underset{N\to\infty}{\longrightarrow} \alpha\mathbb{E}_{\omega,z,A}\left[g_{out}(\omega, \varphi^0_{out}(z,A), \sigma^t) g_{out}(\omega, \varphi^0_{out}(z,A), \sigma^t)^\intercal \right] \end{cases} \tag{35}$$

where $z, \omega \sim \mathcal{N}\left(0, \mathbf{Q}_w^t\right)$ with $\mathbf{Q}^t = \begin{bmatrix} Q & q^t \\ q^t & q^t \end{bmatrix}$.

D. Consistence between replicas and AMP - Bayes optimal case

For the purpose of this section, let's introduce the following non-normalized distributions:

$$\begin{cases}\bar{P}_0(w;T,\Sigma) \equiv P_0(w)\exp\left[-\frac{1}{2}w^\intercal \Sigma w + T^\intercal w\right] \\ f_0(T,\Sigma) \equiv \int_{\mathbb{R}^K} dw \bar{P}_0(w;T,\Sigma) \\ g_0(T,\Sigma) \equiv \mathbb{E}_{\bar{P}_0}[w] = \mathbb{E}_{P_0}[w]\end{cases} \qquad \begin{cases}\bar{P}_{\rm out}(z;\omega,y,V) \equiv e^{-\frac{1}{2}(z-\omega)^\intercal V^{-1}(z-\omega)} P_{\rm out}(y|z) \\ \mathcal{Z}_{out}(\omega,y,V) \equiv \int_{\mathbb{R}^K} dz \bar{P}_{\rm out}(z;\omega,y,V) \\ f_{\rm out}(\omega,y,V) = \int_{\mathbb{R}^K} \mathcal{D}z P_{\rm out}\left(y|V^{1/2}z+\omega\right) = \frac{\det V^{-1/2}}{(2\pi)^{K/2}}\mathcal{Z}_{out} \\ g_{\rm out}(\omega,y,V)) \equiv V^{-1}\mathbb{E}_{\bar{P}_{\rm out}}\left[z-\omega\right] = V^{-1}\mathbb{E}_{P_{\rm out}}\left[z-\omega\right]\end{cases}$$

1. State evolution - AMP

In the Bayes optimal case, using the change of variable $\xi \leftarrow \xi + (\hat{q}^t)^{1/2} W_0$, Eq. 35 becomes:

$$q^{t+1} = \mathbb{E}_\xi\left[f_0\left((\hat{q}^t)^{1/2}\xi,(\hat{q}^t)^{-1}\right) g_0\left((\hat{q}^t)^{1/2}\xi,(\hat{q}^t)^{-1}\right) g_0\left((\hat{q}^t)^{1/2}\xi,(\hat{q}^t)^{-1}\right)^\intercal\right]$$

In addition in the Bayes optimal case, as:

$$\begin{cases}\mathbb{E}_X\left[\omega^t_{\mu\to i}(z_\mu - \omega^t_{\mu\to i})^\intercal\right] = m^t - q^t = 0 \\ \mathbb{E}_X[\omega^t_{\mu\to i}(\omega^t_{\mu\to i})^\intercal] = q^t \\ \mathbb{E}_X\left[(z_\mu^\intercal - \omega^t_{\mu\to i})(z_\mu - \omega^t_{\mu\to i})^\intercal\right] = Q - q^t\,,\end{cases}$$

the multivariate distribution can be written as a product: $\mathcal{N}_{z,\omega}\left(0,\mathbf{Q}^t_w\right) = \mathcal{N}_\omega\left(0,q^t\right)\mathcal{N}_z\left(\omega, Q-q^t\right)$. Hence, using $P_{\rm out}(y|z) = \int dP(A)\delta\left(y-\varphi^0_{\rm out}(z,A)\right)$, Eq. 35 becomes:

$$\begin{aligned}\hat{q}^t &= \alpha\mathbb{E}_{\omega,z,A}\left[g_{out}(\omega,\varphi^0_{out}(z,A),Q-q^t)g_{out}(\omega,\varphi^0_{\rm out}(z,A),Q-q^t)^\intercal\right] \\ &= \alpha\int dy\int d\omega \frac{e^{-\frac{1}{2}\omega^\intercal(q^t)^{-1}\omega}}{(2\pi)^{K/2}\det(q^t)^{1/2}}\int dz P_{\rm out}(y|z)\frac{e^{-\frac{1}{2}(z-\omega)^\intercal(Q-q^t)^{-1}(z-\omega)}}{(2\pi)^{K/2}\det(Q-q^t)^{1/2}} g_{\rm out}(\omega,y,Q-q^t)g_{\rm out}(\omega,y,Q-q^t)^\intercal \\ &= \alpha\int dy\int D\xi\int dz P_{out}(y|z)\frac{e^{-\frac{1}{2}(z-\omega)^\intercal(Q-q^t)^{-1}(z-\omega)}}{(2\pi)^{K/2}\det(Q-q^t)^{1/2}} g_{\rm out}((q^t)^{1/2}\xi,y,Q-q^t)g_{\rm out}((q^t)^{1/2}\xi,y,Q-q^t)^\intercal \\ &= \alpha\mathbb{E}_{y,\xi}\left[f_{\rm out}\left((q^t)^{1/2}\xi,y,Q-q^t\right) g_{\rm out}\left((q^t)^{1/2}\xi,y,Q-q^t\right) g_{\rm out}\left((q^t)^{1/2}\xi,y,Q-q^t\right)^\intercal\right]\end{aligned}$$

Finally to summarize:

$$\begin{cases}q^{t+1} = \mathbb{E}_\xi\left[f_0\left((\hat{q}^t)^{1/2}\xi,(\hat{q}^t)^{-1}\right) g_0\left((\hat{q}^t)^{1/2}\xi,(\hat{q}^t)^{-1}\right) g_0\left((\hat{q}^t)^{1/2}\xi,(\hat{q}^t)^{-1}\right)^\intercal\right] \\ \hat{q}^t = \alpha\mathbb{E}_{y,\xi}\left[f_{\rm out}\left((q^t)^{1/2}\xi,y,Q-q^t\right) g_{\rm out}\left((q^t)^{1/2}\xi,y,Q-q^t\right) g_{\rm out}\left((q^t)^{1/2}\xi,y,Q-q^t\right)^\intercal\right]\end{cases} \tag{36}$$

2. State evolution - Replicas

$$\lim_{N\to\infty} f_n = \text{extr}_{q,\hat{q}}\left\{-\frac{1}{2}\text{Tr}[q\hat{q}] + I_P + \alpha I_C\right\} \quad \text{with} \quad \begin{cases}I_P &\equiv \mathbb{E}_\xi\left[f_0(\hat{q}^{1/2}\xi,\hat{q})\log(f_0(\hat{q}^{1/2}\xi,\hat{q}))\right] \\ I_C &\equiv \mathbb{E}_{\xi,y}\left[f_{\rm out}(q^{1/2}\xi,y,Q-q)\log(f_{\rm out}(q^{1/2}\xi,y,Q-q))\right]\end{cases}$$

Taking the derivatives with respect to q and $\hat{q}$, using an integration by part and the following identities:

$$\begin{cases}\frac{\partial f_{\rm out}}{\partial q} = -\frac{1}{2}q^{-1}e^{\frac{1}{2}\xi^\intercal\xi}\partial_\xi\left[e^{-\frac{1}{2}\xi^\intercal\xi}\partial_\xi f_{\rm out}\right] \\ \frac{\partial f_0}{\partial \hat{q}} = -\frac{1}{2}\hat{q}^{-1}e^{\frac{1}{2}\xi^\intercal\xi}\partial_\xi\left[e^{-\frac{1}{2}\xi^\intercal\xi}\partial_\xi f_0\right]\,,\end{cases}$$

the state evolution equations read:

$$\begin{cases} q = 2\frac{\partial I_P}{\partial \hat{q}} \\ \hat{q} = 2\alpha \frac{\partial I_C}{\partial q} \end{cases} \quad \text{with} \quad \begin{cases} \frac{\partial I_P}{\partial \hat{q}} = \frac{1}{2}\mathbb{E}_\xi \left[f_0(\hat{q}^{1/2}\xi, \hat{q}) g_0(\hat{q}^{1/2}\xi, \hat{q}) g_0(\hat{q}^{1/2}\xi, \hat{q})^\intercal \right] \\ \frac{\partial I_C}{\partial q} = \frac{1}{2}\mathbb{E}_{y,\xi} \left[f_{\mathrm{out}}(q^{1/2}\xi, y, Q-q) g_{\mathrm{out}}(q^{1/2}\xi, y, Q-q) g_{\mathrm{out}}(q^{1/2}\xi, y, Q-q)^\intercal \right] \end{cases}$$

that simplify allows to recover Eq.36 without time indices:

$$\begin{cases} q = \mathbb{E}_\xi \left[f_0(\hat{q}^{1/2}\xi, \hat{q}) g_0(\hat{q}^{1/2}\xi, \hat{q}) g_0(\hat{q}^{1/2}\xi, \hat{q})^\intercal \right] \\ \hat{q} = \alpha \mathbb{E}_{y,\xi} \left[f_{\mathrm{out}}(q^{1/2}\xi, y, Q-q) g_{\mathrm{out}}(q^{1/2}\xi, y, Q-q) g_{\mathrm{out}}(q^{1/2}\xi, y, Q-q)^\intercal \right] \end{cases}$$

E. RS assumption

a. Trace term

$$\sum_{ab}\hat{Q}_{ab}Q_{ab}=Q^0\hat{Q}^0+2nm\hat{m}+nQ_w\hat{Q}+n(n-1)\underline{\underline{q}}\hat{q} \tag{37}$$

b. Integral $\hat{\mathcal{I}}_w$

$$\hat{\mathcal{I}}_w(\underline{\underline{\hat{\mathbf{Q}}}}^w)=\int\prod_{a=0}^n dw^a P_{w^a}(w^a)\exp\left(-\sum_{(a,b)}\hat{Q}_{ab}w^a w^b\right)$$
$$=\int\prod_{a=0}^n dw^a P_{w^a}(w^a)\exp\left(-\left[w^0\hat{Q}^0w^0+2\sum_{a=1}^n w^0\hat{m}w^a+\sum_{a=1}^n w^a\hat{Q}w^a+\sum_{a,b=1,a\neq b}^N (w^b)^\intercal\hat{q}w^a\right]\right)$$
$$=\int\prod_{a=0}^n dw^a P_{w^a}(w^a)e^{-\left[(w^0)^\intercal\hat{Q}^0w^0+2\sum_{a=1}^n(w^0)^\intercal\hat{m}w^a+\sum_{a=1}^n(w^a)^\intercal\hat{Q}w^a\right]}\exp\left(-\sum_{a,b=1}^N w^b\hat{q}w^a+\sum_{a=1}^n w^a\hat{q}w^a\right)$$

Finally to decouple replicas we use a Hubbard Stratanovich (Gaussian transformation), where $D\xi=\frac{e^{-\xi^2/2}}{\sqrt{2\pi}}$:

$$\int D\xi e^{x\sqrt{q}\xi}=e^{qx^2}$$

Finally:

$$\hat{\mathcal{I}}_w(\underline{\underline{\hat{\mathbf{Q}}}})=\int D\xi\int\prod_{a=0}^n dw^a P_{w^a}(w^a)e^{-\left[(w^0)^\intercal\hat{Q}^0w^0+2\sum_{a=1}^n(w^0)^\intercal\hat{m}w^a+\sum_{a=1}^n(w^a)^\intercal\hat{Q}w^a\right]}\exp\left(-\sum_{a=1}^N\xi^\intercal\hat{q}^{1/2}w^a+\sum_{a=1}^n w^a\hat{q}w^a\right)$$
$$=\int_{\mathbb{R}}D\xi\int_{\mathbb{R}}dw^0P_{w^0}(w^0)e^{-(w^0)^\intercal\hat{Q}^0_w w^0}\left[\int_{\mathbb{R}}dwP_w(w)\exp\left(-\left[2w^0\hat{m}w+w\left(\hat{Q}+\hat{q}\right)w-\xi\hat{q}^{1/2}w\right]\right)\right]^n$$

c. Integral $\mathcal{I}_z$

$$\hat{\mathcal{I}}_z(\underline{\underline{\mathbf{Q}}};\Delta)\equiv\int dy\int_{\mathbb{R}^{n+1}}\prod_{a=0}^n dz^a P^a_{out}\left(y|\varphi^a_{out}\left(z^a;\Delta\right)\right)\times\exp\left(-\frac{1}{2}\sum_{ab}z^az^b(\Sigma^{ab})^{-1}-\frac{1}{2}\log(\det\underline{\underline{\boldsymbol{\Sigma}}})-\frac{(n+1)}{2}\log(2\pi)\right)$$

To explicit the integral we need to express $\underline{\underline{\boldsymbol{\Sigma}}}^{-1}$:

$$\underline{\underline{\boldsymbol{\Sigma}}}^{-1}=\begin{bmatrix}\Sigma^{-1}_{00}&\Sigma^{-1}_{01}&\Sigma^{-1}_{01}&\Sigma^{-1}_{01}\\\Sigma^{-1}_{01}&\Sigma^{-1}_{11}&\Sigma^{-1}_{12}&\Sigma^{-1}_{12}\\\Sigma^{-1}_{01}&\Sigma^{-1}_{12}&\Sigma^{-1}_{11}&\Sigma^{-1}_{12}\\\Sigma^{-1}_{01}&\Sigma^{-1}_{12}&\Sigma^{-1}_{12}&\Sigma^{-1}_{11}\end{bmatrix} \tag{38}$$

with

$$\begin{cases} \Sigma_{00}^{-1} & = \left(Q^0 - nm(Q+(n-1)q)^{-1}m\right)^{-1} \\ \Sigma_{01}^{-1} & = -\left(Q^0 - nm(Q+(n-1)q)^{-1}m\right)^{-1} m(Q+(n-1)q)^{-1} \\ \Sigma_{11}^{-1} & = (Q-q)^{-1} - (Q+(n-1)q)^{-1}q(Q-q)^{-1} \\ & +(Q+(n-1)q)^{-1}m\left(Q^0 - nm(Q+(n-1)q)^{-1}m\right)^{-1} m(Q_w+(n-1)q)^{-1} \\ \Sigma_{12}^{-1} & = -(Q+(n-1)q)^{-1}q(Q-q)^{-1} \\ & +(Q+(n-1)q)^{-1}m\left(Q - nm(Q+(n-1)q)^{-1}m\right)^{-1} m(Q+(n-1)q)^{-1} \end{cases}$$

and its determinant:

$$\det \underline{\underline{\mathbf{\Sigma}}} = \det\left(Q-q\right)^{n-1} \det\left(Q+(n-1)q\right) \det\left(Q^0 - nm(Q+(n-1)q)^{-1}m\right)$$

In the same way than above, introducing the gaussian transformation:

$$\exp\left(-\frac{1}{2}\sum_{ab} z^a z^b (\Sigma_{ab})^{-1}\right) = \exp\left[-\frac{1}{2}\left(z^0\Sigma_{00}^{-1}z^0 + 2\sum_{a=1}^{n}(z^0)^\intercal\Sigma_{01}^{-1}z^a + \sum_{a=1}^{n}(z^a)^\intercal\Sigma_{11}^{-1}z^a + \sum_{a,b=1,a\neq b}^{N} z^b\Sigma_{12}^{-1}z^a\right)\right]$$
$$= \int D\xi \exp\left(-\frac{1}{2}\left(z^0\Sigma_{00}^{-1}z^0 + 2\sum_{a=1}^{n} z^0\Sigma_{01}^{-1}z^a + \sum_{a=1}^{n} z^a\Sigma_{11}^{-1}z^a\right)\right) \exp\left(-\sum_{a=1}^{n}\xi^\intercal\Sigma_{12}^{-1/2}z^a + \frac{1}{2}\sum_{a=1}^{n} z^a\Sigma_{12}^{-1}z^a\right)$$
$$= \int D\xi \exp\left(-\frac{1}{2}\left((z^0)^\intercal\Sigma_{00}^{-1}z^0 + 2\sum_{a=1}^{n} z^0\Sigma_{01}^{-1}z^a + \sum_{a=1}^{n} z^a\Sigma_{11}^{-1}z^a + \sum_{a=1}^{n} 2\xi\Sigma_{12}^{-1/2}z^a - \sum_{a=1}^{n} z^a\Sigma_{12}^{-1}z^a\right)\right)$$

Finally the integral can be put under the following form:

$$\hat{\mathcal{I}}_z(\underline{\underline{\mathbf{Q}}};\Delta) = \int D\xi \int dy e^{-\frac{1}{2}\log(\det\underline{\underline{\mathbf{\Sigma}}}) - \frac{(n+1)}{2}\log(2\pi)} \int dz^0 P_{out}^0\left(y|\varphi_{out}^0\left(z^0;\Delta^0\right)\right) e^{-\frac{1}{2}\left(z^0\Sigma_{00}^{-1}z^0\right)}$$
$$\left[\int dz P_{out}\left(y|\varphi_{out}\left(z;\Delta\right)\right) e^{-z^0\Sigma_{01}^{-1}z - \frac{1}{2}z\left(\Sigma_{11}^{-1} - \Sigma_{12}^{-1}\right)z - \xi\Sigma_{12}^{-1/2}z}\right]^n$$

F. Check $n \to 0$

In fact in this limit:

$$\begin{cases} \mathcal{H}(\underline{\underline{\mathbf{Q}}}^w, \underline{\underline{\hat{\mathbf{Q}}}}^w) & \equiv \mathrm{Tr} Q^0 \underline{\underline{\hat{Q}}}^0_w + \log\left(\hat{\mathcal{I}}^0_w\right) + \alpha_2 \log\left(\hat{\mathcal{I}}^0_z\right) \\ \hat{\mathcal{I}}^0_w & \equiv \int d\underline{w}^0 P_{w^0}(\underline{w}^0) \exp\left(-(\underline{w}^0)^\intercal \underline{\underline{\hat{Q}}}^0_w \underline{w}^0\right) \\ \hat{\mathcal{I}}^0_z & \equiv \int dy \int d\underline{z}^0 P^0_{out}\left(\underline{\tilde{u}} | \varphi^0_{out}\left(\underline{z}^0, \Delta^0_2; \underline{\underline{\Gamma}}^0_v, \Delta^0_v\right)\right) \exp\left(-\frac{1}{2}(\underline{z}^0)^\intercal \Sigma^{-1}_{00} \underline{z}^0 - \frac{1}{2}\log(\lim_{n\to 0} \det \underline{\underline{\mathbf{\Sigma}}}) - \frac{K}{2}\log(2\pi)\right) \\ & = \int d\underline{z}^0 \exp\left(-\frac{1}{2}(\underline{z}^0)^\intercal \Sigma^{-1}_{00} \underline{z}^0 - \frac{1}{2}\log(\lim_{n\to 0} \det \underline{\underline{\mathbf{\Sigma}}}) - \frac{K}{2}\log(2\pi)\right) \end{cases} \tag{39}$$

with $\Sigma^{-1}_{00} = (Q^0)^{-1}$ and $\lim_{n\to 0} \det \underline{\underline{\mathbf{\Sigma}}} = \det(Q^0)$
Hence, putting things together:
Doing so it leads to :

$$\begin{aligned} \mathcal{H}(\underline{\underline{\mathbf{Q}}}^w, \underline{\underline{\hat{\mathbf{Q}}}}^w) = \mathrm{Tr} Q^0 \underline{\underline{\hat{Q}}}^0_w + \int d\underline{w}^0 P_{w^0}(\underline{w}^0) \exp\left(-(\underline{w}^0)^\intercal \underline{\underline{\hat{Q}}}^0_w \underline{w}^0\right) = 0 \\ \iff \begin{cases} \underline{\underline{\hat{Q}}}^0_w = 0 \\ Q^0 = \mathbb{E}_{P^0}\left[\underline{w}^0 (\underline{w}^0)^\intercal\right] \end{cases} \end{aligned}$$

G. $n \to 0$ limit

- Integral $\mathcal{I}_w$

$$\begin{aligned} &\mathcal{I}_w\left(\underline{\underline{\hat{Q}}}^0_w, \underline{\underline{\hat{Q}}}_w, \hat{m}, \underline{\underline{\hat{q}}}_w\right) := \lim_{n\to 0} \frac{1}{N} \log\left(\hat{\mathcal{I}}_w(\underline{\underline{\hat{\mathbf{Q}}}}^w)\right) = \\ &= \lim_{n\to 0} \frac{1}{N} \log\left(\int D\xi \int d\underline{w}^0 P_{w^0}(\underline{w}^0) e^{-(\underline{w}^0)^\intercal \underline{\underline{\hat{Q}}}^0_w \underline{w}^0} \left[\int_{\mathbb{R}^K} d\underline{w} P_w(\underline{w}) \exp\left(-\left[2(\underline{w}^0)^\intercal \hat{m} \underline{w} + (\underline{w})^\intercal \left(\underline{\underline{\hat{Q}}}_w + \underline{\underline{\hat{q}}}_w\right) \underline{w} - \xi^\intercal \underline{\underline{\hat{q}}}^{1/2}_w \underline{w}\right]\right)\right]^N\right) \\ &= \int D\xi \int d\underline{w}^0 P_{w^0}(\underline{w}^0) e^{-(\underline{w}^0)^\intercal \underline{\underline{\hat{Q}}}^0_w \underline{w}^0} \log\left(\int_{\mathbb{R}^K} d\underline{w} P_w(\underline{w}) \exp\left(-2(\underline{w}^0)^\intercal \hat{m} \underline{w} - (\underline{w})^\intercal \left(\underline{\underline{\hat{Q}}}_w + \underline{\underline{\hat{q}}}_w\right) \underline{w} + \xi^\intercal \underline{\underline{\hat{q}}}^{1/2}_w \underline{w}\right)\right) \end{aligned}$$

- Integral $\mathcal{I}_z$

We first take the limite of the determinant: $\det \underline{\underline{\mathbf{\Sigma}}} \xrightarrow{n\to 0} \det\left(\underline{\underline{Q}}^0_w\right)^{-1}$, then of the matrix terms:

$$\begin{cases} \Sigma^{-1}_{00} & \xrightarrow{n\to 0} \left(\underline{\underline{Q}}^0_w\right)^{-1} \\ \Sigma^{-1}_{01} & \xrightarrow{n\to 0} -\left(\underline{\underline{Q}}^0_w\right)^{-1} m(\underline{\underline{Q}}_w - \underline{\underline{q}}_w)^{-1} \\ \Sigma^{-1}_{11} & \xrightarrow{n\to 0} (\underline{\underline{Q}}_w - \underline{\underline{q}}_w)^{-1} \left(\underline{\underline{\mathbb{1}}} - \underline{\underline{q}}_w + m^\intercal \left(\underline{\underline{Q}}^0_w\right)^{-1} m\right) (\underline{\underline{Q}}_w - \underline{\underline{q}}_w)^{-1} \\ \Sigma^{-1}_{12} & \xrightarrow{n\to 0} (\underline{\underline{Q}}_w - \underline{\underline{q}}_w)^{-1} \left(-\underline{\underline{q}}_w + m^\intercal \left(\underline{\underline{Q}}^0_w\right)^{-1} m\right) (\underline{\underline{Q}}_w - \underline{\underline{q}}_w)^{-1} \equiv -\Omega \end{cases}$$

a. Detail of the $n \to 0$ limit

Let's write:

$$\hat{\mathcal{I}}_z(\underline{\underline{\mathbf{Q}}}^w;\Delta_2;\underline{\underline{\Gamma}}_v,\Delta_v) = \int D\xi \int dy \int \frac{d\underline{z}^0}{(2\pi)^{K/2}} P^0_{out}\left(\underline{\underline{\tilde{u}}}|\varphi^0_{out}\left(\underline{z}^0,\Delta^0_2;\underline{\underline{\Gamma}}^0_v,\Delta^0_v\right)\right) e^{-\frac{1}{2}(\underline{z}^0)^\intercal \left(\underline{\underline{Q}}^0_w\right)^{-1}\underline{z}^0 - \frac{1}{2}\log\det\left(\underline{\underline{Q}}^0_w\right)^{-1}}$$
$$\left[\int \frac{d\underline{z}}{(2\pi)^{K/2}} P_{out}\left(\underline{\underline{\tilde{u}}}|\varphi_{out}\left(\underline{z},\Delta_2;\underline{\underline{\Gamma}}_v,\Delta_v\right)\right) e^{(\underline{z}^0)^\intercal\left(\underline{\underline{Q}}^0_w\right)^{-1} m(\underline{\underline{Q}}_w - \underline{\underline{q}}_w)^{-1}\underline{z} - \frac{1}{2}\underline{z}^\intercal(\underline{\underline{Q}}_w - \underline{\underline{q}}_w)^{-1}\underline{z} + \xi^\intercal\underline{\underline{\Omega}}^{1/2}\underline{z}}\right]^N$$
$$\equiv \int dy \int D\xi J_0(y,\xi,n) J_1(y,\xi,n)^N$$

• *Method 1:*
The expansion to the first order can be written as, where the last line has to be detailed:

$$\begin{aligned}
\hat{\mathcal{I}}_z(\underline{\underline{\mathbf{Q}}}^w;\Delta_2;\underline{\underline{\Gamma}}_v,\Delta_v) &= \int dy \int D\xi J_0(y,\xi,n) J_1(y,\xi,n)^N \\
&\simeq \int dy \int D\xi J_0(y,\xi,0) + n\int dy\int D\xi\partial_n J_0(y,\xi,0) + n\int dy \int D\xi J_0(y,\xi,0)\log\left(J_1(y,\xi,0)\right) \\
&= n\int dy\int D\xi J_0(y,\xi,0)\log\left(J_1(y,\xi,0)\right)
\end{aligned}$$

Hence:

$$\hat{\mathcal{I}}_z(\underline{\underline{\mathbf{Q}}}^w;\Delta_2;\underline{\underline{\Gamma}}_v,\Delta_v) = \int D\xi \int dy \int \frac{d\underline{z}^0}{(2\pi)^{K/2}} P^0_{out}\left(\underline{\underline{\tilde{u}}}|\varphi^0_{out}\left(\underline{z}^0,\Delta^0_2;\underline{\underline{\Gamma}}^0_v,\Delta^0_v\right)\right) e^{-\frac{1}{2}(\underline{z}^0)^\intercal \left(\underline{\underline{Q}}^0_w\right)^{-1}\underline{z}^0 - \frac{1}{2}\log\det\left(\underline{\underline{Q}}^0_w\right)^{-1}}$$
$$\left[\int \frac{d\underline{z}}{(2\pi)^{K/2}} P_{out}\left(\underline{\underline{\tilde{u}}}|\varphi_{out}\left(\underline{z},\Delta_2;\underline{\underline{\Gamma}}_v,\Delta_v\right)\right) e^{(\underline{z}^0)^\intercal\left(\underline{\underline{Q}}^0_w\right)^{-1} m(\underline{\underline{Q}}_w - \underline{\underline{q}}_w)^{-1}\underline{z} - \frac{1}{2}\underline{z}^\intercal(\underline{\underline{Q}}_w - \underline{\underline{q}}_w)^{-1}\underline{z} + \xi^\intercal\underline{\underline{\Omega}}^{1/2}\underline{z}}\right]^N$$

Changing the variable $\underline{z}$ and shifting $\xi' = \xi + \underline{\underline{\Omega}}^{-1/2}(\underline{\underline{Q}}_w - \underline{\underline{q}}_w)^{-1}\underline{\underline{m}}^\intercal(\underline{\underline{Q}}^0)^{-1}\underline{z}^0$,

$$\hat{\mathcal{I}}_z(\underline{\underline{\mathbf{Q}}}^w;\Delta_2;\underline{\underline{\Gamma}}_v,\Delta_v) = \int D\xi \int dy \int \frac{d\underline{z}^0}{(2\pi)^{K/2}} P^0_{out}\left(\underline{\underline{\tilde{u}}}|\varphi^0_{out}\left(\underline{z}^0,\Delta^0_2;\underline{\underline{\Gamma}}^0_v,\Delta^0_v\right)\right) e^{-\frac{1}{2}(\underline{z}^0)^\intercal\underline{\underline{N}}\underline{z}^0 + \xi^\intercal\underline{\underline{M}}\underline{z}^0 - \frac{1}{2}\log\det\left(\underline{\underline{Q}}^0_w\right)^{-1}}$$
$$\left[\int D\underline{z} P_{out}\left(\underline{\underline{\tilde{u}}}|\varphi_{out}\left(\underline{z},\Delta_2;\underline{\underline{\Gamma}}_v,\Delta_v\right)\right) e^{\xi^\intercal\underline{\underline{\Omega}}^{1/2}(\underline{\underline{Q}} - \underline{\underline{q}})^{1/2}\underline{z}}\right]^N$$

where:

$$\begin{cases}
\underline{\underline{M}} = \underline{\underline{\Omega}}^{-1/2}(\underline{\underline{Q}}_w - \underline{\underline{q}}_w)^{-1}\underline{\underline{m}}^\intercal(\underline{\underline{Q}}^0_w)^{-1} \\
\underline{\underline{N}} = (\underline{\underline{Q}}^0_w)^{-1} + \underline{\underline{M}}\underline{\underline{M}}^\intercal = \dots = (\underline{\underline{Q}}^0_w - \underline{\underline{m}}\underline{\underline{q}}^{-1}\underline{\underline{m}}^\intercal)^{-1}
\end{cases}$$

We rescale $\underline{z}^0$ by $\underline{\underline{N}}^{-1/2}$

$$\hat{\mathcal{I}}_z(\underline{\underline{\mathbf{Q}}}^w;\Delta_2;\underline{\underline{\Gamma}}_v,\Delta_v) = \int D\xi \int dy \int D\underline{z}^0 P^0_{out}\left(y|\tilde{\varphi}^0_{out}\left((\underline{\underline{Q}}^0_w - \underline{\underline{m}}\underline{\underline{q}}^{-1}\underline{\underline{m}}^\intercal)^{1/2}\underline{z}^0,\Delta^0_2;\underline{\underline{\Gamma}}^0_v,\Delta^0_v\right)\right)$$
$$e^{\xi^\intercal\underline{\underline{M}}\underline{\underline{N}}^{-1/2}\underline{z}^0 - \frac{1}{2}\log\det\left(\underline{\underline{Q}}^0_w\right)^{-1} + \frac{1}{2}\log\det\left(\underline{\underline{Q}}^0_w - \underline{\underline{m}}\underline{\underline{q}}^{-1}\underline{\underline{m}}^\intercal\right)}$$
$$\left[\int D\underline{z} P_{out}\left(y|\varphi_{out}\left((\underline{\underline{Q}}_w - \underline{\underline{q}})^{1/2}\underline{z},\Delta_2;\underline{\underline{\Gamma}}_v,\Delta_v\right)\right) e^{\xi^\intercal\underline{\underline{\Omega}}^{1/2}(\underline{\underline{Q}} - \underline{\underline{q}})^{1/2}\underline{z}}\right]^N$$

where:

$$\begin{cases} \underline{\underline{MN}}^{-1/2} = \underline{\underline{\Omega}}^{1/2}(\underline{\underline{Q}}_w - \underline{\underline{q}})^{1/2}((\underline{\underline{Q}}_w - \underline{\underline{q}})^{1/2}\underline{\underline{q}}_w^{-1}\underline{\underline{m}}^\intercal)(\underline{\underline{Q}}^0_w - \underline{\underline{m}}\underline{\underline{q}}^{-1}\underline{\underline{m}}^\intercal)^{-1/2} = \underline{\underline{\Omega}}^{1/2}(\underline{\underline{Q}}_w - \underline{\underline{q}})^{1/2}\underline{\underline{K}} \\ \underline{\underline{K}} = ((\underline{\underline{Q}}_w - \underline{\underline{q}})^{1/2}\underline{\underline{q}}_w^{-1}\underline{\underline{m}}^\intercal)(\underline{\underline{Q}}^0_w - \underline{\underline{m}}\underline{\underline{q}}^{-1}\underline{\underline{m}}^\intercal)^{-1/2} \end{cases}$$

Shiftting $\underline{z}$: $\underline{z} \leftarrow \underline{z} + \underline{\underline{\Omega}}^{1/2}(\underline{\underline{Q}} - \underline{\underline{q}})^{1/2}\xi$ and $\underline{z}_0 \leftarrow \underline{z}_0 - \Omega^{1/2}(\underline{\underline{Q}}_w - \underline{\underline{q}}_w)^{1/2}\underline{\underline{K}}\xi$ and taking $n \to 0$, we obtain:

$$\hat{\mathcal{I}}_z(\underline{\underline{\mathbf{Q}}}^w;\Delta_2;\underline{\underline{\Gamma}}_v,\Delta_v) = \int \frac{d\xi}{(2\pi)^{K/2}} \int dy \int D\underline{z}^0 P^0_{out}\left(y|\tilde{\varphi}^0_{out}\left((\underline{\underline{Q}}^0_w - m\underline{\underline{q}}_w^{-1}m^\intercal)^{1/2}\underline{z}^0 + m\underline{\underline{q}}_w^{-1}(\underline{\underline{Q}}_w - \underline{\underline{q}}_w)\Omega^{1/2}\xi, \Delta^0_2; \underline{\underline{\Gamma}}^0_v, \Delta^0_v\right)\right)$$
$$e^{-\frac{1}{2}\xi^\intercal\left[\underline{\underline{\mathbb{1}}} - \underline{\underline{\Omega}}^{1/2}(\underline{\underline{Q}}_w - \underline{\underline{q}}_w)^{1/2}\underline{\underline{KK}}^\intercal(\underline{\underline{Q}}_w - \underline{\underline{q}}_w)^{1/2}\Omega^{1/2}\xi\right] - \frac{1}{2}\log\det\left(\underline{\underline{Q}}^0_w\right)^{-1} + \frac{1}{2}\log\det\left(\underline{\underline{Q}}^0_w - m\underline{\underline{q}}_w^{-1}\underline{\underline{m}}^\intercal\right)}$$
$$\left[\int D\underline{z}P_{out}\left(y|\varphi_{out}\left((\underline{\underline{Q}}_w - \underline{\underline{q}}_w)^{1/2}\underline{z} + (\underline{\underline{Q}} - \underline{\underline{q}}_w)^{1/2}\underline{\underline{\Omega}}^{1/2}\xi, \Delta_2; \underline{\underline{\Gamma}}_v, \Delta_v\right)\right)\right]^N$$

However:

$$\left[\underline{\underline{\mathbb{1}}} - \underline{\underline{\Omega}}^{1/2}(\underline{\underline{Q}}_w - \underline{\underline{q}}_w)^{1/2}\underline{\underline{KK}}^\intercal(\underline{\underline{Q}}_w - \underline{\underline{q}}_w)^{1/2}\Omega^{1/2}\xi\right] = \underline{\underline{\Omega}}^{1/2}(\underline{\underline{Q}}_w - \underline{\underline{q}}_w)\underline{\underline{q}}_w^{-1}(\underline{\underline{Q}}_w - \underline{\underline{q}}_w)\underline{\underline{\Omega}}^{1/2}$$

Doing the change of variable: $\xi \leftarrow \underline{\underline{q}}_w^{-1/2}(\underline{\underline{Q}}_w - \underline{\underline{q}}_w)\underline{\underline{\Omega}}^{1/2}\xi$

$$\hat{\mathcal{I}}_z(\underline{\underline{\mathbf{Q}}}^w;\Delta_2;\underline{\underline{\Gamma}}_v,\Delta_v) = \int D\xi \int dy \int D\underline{z}^0 P^0_{out}\left(y|\tilde{\varphi}^0_{out}\left((\underline{\underline{Q}}^0_w - m\underline{\underline{q}}_w^{-1}m^\intercal)^{1/2}\underline{z}^0 + m\underline{\underline{q}}_w^{-1/2}\xi, \Delta^0_2; \underline{\underline{\Gamma}}^0_v, \Delta^0_v\right)\right)$$
$$e^{-\frac{1}{2}\log\det\left(\underline{\underline{Q}}^0_w\right)^{-1} + \frac{1}{2}\log\det\left(\underline{\underline{Q}}^0_w - m\underline{\underline{q}}_w^{-1}\underline{\underline{m}}^\intercal\right) - \frac{1}{2}\log\det\left(\underline{\underline{q}}_w^{-1}(\underline{\underline{Q}}_w - \underline{\underline{q}}_w)^2\underline{\underline{\Omega}}\right)}$$
$$\left[\int D\underline{z}P_{out}\left(y|\varphi_{out}\left((\underline{\underline{Q}}_w - \underline{\underline{q}}_w)^{1/2}\underline{z} + \underline{\underline{q}}_w^{1/2}\xi, \Delta_2; \underline{\underline{\Gamma}}_v, \Delta_v\right)\right)\right]^N$$

The exponential with determinants vanishes because:

$$\begin{aligned} &-\frac{1}{2}\log\det\left(\underline{\underline{Q}}^0_w\right)^{-1} + \frac{1}{2}\log\det\left(\underline{\underline{Q}}^0_w - m\underline{\underline{q}}_w^{-1}\underline{\underline{m}}^\intercal\right) - \frac{1}{2}\log\det\left(\underline{\underline{q}}_w^{-1}(\underline{\underline{Q}}_w - \underline{\underline{q}}_w)^2\underline{\underline{\Omega}}\right) \\ &= -\frac{1}{2}\log\det\left(\underline{\underline{Q}}^0_w\right)^{-1} + \frac{1}{2}\log\det\left(\underline{\underline{Q}}^0_w - m\underline{\underline{q}}_w^{-1}\underline{\underline{m}}^\intercal\right) - \frac{1}{2}\log\det\left(\underline{\underline{q}}_w^{-1}\left(-\underline{\underline{q}}_w + m^\intercal\left(\underline{\underline{Q}}^0_w\right)^{-1}m\right)\right) \\ &= 0 \end{aligned}$$

In the end:

$$\hat{\mathcal{I}}_z(\underline{\underline{\mathbf{Q}}}^w;\Delta_2;\underline{\underline{\Gamma}}_v,\Delta_v) = \int D\xi \int dy \int D\underline{z}^0 P^0_{out}\left(y|\tilde{\varphi}^0_{out}\left((\underline{\underline{Q}}^0_w - m\underline{\underline{q}}_w^{-1}m^\intercal)^{1/2}\underline{z}^0 + m\underline{\underline{q}}_w^{-1/2}\xi, \Delta^0_2; \underline{\underline{\Gamma}}^0_v, \Delta^0_v\right)\right)$$
$$\left[\int D\underline{z}P_{out}\left(y|\varphi_{out}\left((\underline{\underline{Q}}_w - \underline{\underline{q}}_w)^{1/2}\underline{z} + \underline{\underline{q}}_w^{1/2}\xi, \Delta_2; \underline{\underline{\Gamma}}_v, \Delta_v\right)\right)\right]^N$$

and finally:

$$\begin{aligned} &\mathcal{I}_z\left(Q^0, \underline{\underline{Q}}_w, m, \underline{\underline{q}}_w; \Delta_2; \underline{\underline{\Gamma}}_v, \Delta_v\right) := \lim_{n\to 0}\frac{1}{N}\log\left(\hat{\mathcal{I}}_z(\underline{\underline{\mathbf{Q}}}^w;\Delta_2;\underline{\underline{\Gamma}}_v,\Delta_v)\right) \\ &= \int dy \int D\xi \int D\underline{z}^0 P^0_{out}\left(y|\varphi^0_{out}\left(\left(Q^0 - m\underline{\underline{q}}_w^{-1}m^\intercal\right)^{1/2}\underline{z}^0 + m\underline{\underline{q}}_w^{-1/2}\xi, \Delta^0_2; \underline{\underline{\Gamma}}^0_v, \Delta^0_v\right)\right) \\ &\times \log\left(\int D\underline{z}P_{out}\left(y|\varphi_{out}\left(\left(\underline{\underline{Q}}_w - \underline{\underline{q}}_w\right)^{1/2}\underline{z} + \underline{\underline{q}}_w^{1/2}\xi, \Delta_2; \underline{\underline{\Gamma}}_v, \Delta_v\right)\right)\right) \end{aligned}$$

- *Method 2:*

Cf Bruno's notes: put the determinant term outside and expand the product of $\det(n) \times J_0(n)J_1^N$

H. Non Bayes optimal

1. Non-Bayes optimal setting

$$
\begin{aligned}
\Phi =& \frac{1}{N}\mathbb{E}_{\underline{y},\underline{w}^0,\underline{\underline{X}}}\left[\log \mathcal{Z}(\underline{y};\underline{\underline{X}})\right] \\
&= \mathbf{extr}_{Q,\hat{Q},q,\hat{q},m,\hat{m}}\left\{-m\hat{m} + Q\hat{Q} + \frac{1}{2}q\hat{q} + \mathcal{I}_w\left(\hat{Q}^0,\hat{Q},\hat{m},\hat{q}\right) + \alpha\mathcal{I}_z\left(Q^0,Q,m,q;\Delta^0,\Delta\right)\right\}
\end{aligned}
$$

a. $\mathcal{I}_z$

$$
\begin{aligned}
&\mathcal{I}_z\left(Q^0,Q,m,q;\Delta^0,\Delta\right) \\
&= \int dy \int D\xi \int Dz^0 P^0_{out}\left(y|\varphi^0_{out}\left(\left(Q^0 - mq^{-1}m^\intercal\right)^{1/2} z^0 + mq^{-1/2}\xi;\Delta^0\right)\right) \\
&\times \log\left(\int Dz P_{out}\left(y|\varphi_{out}\left((Q-q)^{1/2}\underline{z} + q^{1/2}\xi;\Delta\right)\right)\right) \\
&= \mathbb{E}_{y,\xi}\left[f_{out}\left(y,\omega^0[\xi],V^0;\Delta^0\right)\log f_{out}\left(y,\omega[\xi],V;\Delta\right)\right]
\end{aligned}
$$

with

$$
\begin{cases}\omega^0[\xi] = mq^{-1/2}\xi \\ V^0 = \left(Q^0 - mq^{-1}m\right)\end{cases} \quad\text{and}\quad \begin{cases}\omega[\xi] = q^{1/2}\xi \\ V = (Q-q)\end{cases} \tag{40}
$$

b. $\mathcal{I}_w$

Using the change of variable $\xi' = \xi - (\hat{q})^{-1/2}\,\hat{m}w^0$:

$$
\begin{aligned}
&\mathcal{I}_w\left(\hat{Q}^0,\hat{Q},\hat{m},\hat{q}\right) = \int D\xi \int dw^0 P_{w^0}(w^0) \\
&\times \log\left(\int dw P_w(w)\exp\left(-w^0\hat{m}w - \frac{1}{2}w\left(\hat{Q}+\hat{q}\right)w + \xi^\intercal\hat{q}^{1/2}w\right)\right) \\
&= \mathbb{E}_\xi\left[\int dw^0 P_{w^0}(w^0) e^{-\frac{1}{2}(w^0)^\intercal\left(\hat{Q}^0+\hat{m}\hat{q}^{-1}\hat{m}\right)w^0 + \xi^\intercal\hat{q}^{-1/2}\hat{m}w^0}\log\left(\int dw P_w(w) e^{\frac{1}{2}w\left(\hat{Q}+\hat{q}\right)w + \xi\hat{q}^{1/2}w}\right)\right] \\
&= \mathbb{E}_\xi\left[f^w_0\left(\lambda^0[\xi],\sigma^0\right)\log f^w_0\left(\lambda[\xi],\sigma\right)\right]
\end{aligned}
$$

with

$$
\begin{cases}\left(\sigma^0\right)^{-1}\lambda^0[\xi] = \hat{m}\hat{q}^{-1/2}\xi \\ \sigma^0 = \left(\hat{Q}^0 + \hat{m}\hat{q}^{-1}\hat{m}\right)^{-1}\end{cases} \quad\text{and}\quad \begin{cases}\sigma^{-1}\lambda[\xi] = \hat{q}^{1/2}\xi \\ \sigma = \left(\hat{Q}+\hat{q}\right)^{-1}\end{cases}
$$

V. References

[1] M. Mezard, G. Parisi, and M. A. Virasoro, *Spin Glass Theory and Beyond* (1987).

[2] J. Barbier, F. Krzakala, N. Macris, L. Miolane, and L. Zdeborová, pp. 1–59 (2017), 1708.03395, URL http://arxiv.org/abs/1708.03395.

[3] F. Krzakala, M. Mézard, F. Sausset, Y. Sun, and L. Zdeborová, Journal of Statistical Mechanics: Theory and Experiment **2012**, 1 (2012), ISSN 17425468, 1206.3953.

[4] B. Aubin, A. Maillard, J. Barbier, F. Krzakala, N. Macris, and L. Zdeborová, pp. 1–44 (2018), 1806.05451, URL http://arxiv.org/abs/1806.05451.

[5] F. Krzakala, M. Mézard, F. Sausset, Y. Sun, and L. Zdeborová, pp. 1–21 (2011), ISSN 2160-3308, 1109.4424, URL http://arxiv.org/abs/1109.4424{%}0Ahttp://dx.doi.org/10.1103/PhysRevX.2.021005.