Statistical Physics of Computation - Mock exam

Vittorio Erba, Emanuele Troiani

Phys-512 – 2025/26

Instructions:

- Duration of the exam: 3 hours.
- Allowed material: 1 page recto-verso of personal notes, paper, material for writing.
- Questions can be solved in any order. They *are not* ordered by difficulty.
- The total number of points is 36.

NOTICE:

- The program last year was different, so the difficulty of the questions is not exactly the same.
- Most questions allow for at least a partial answer (but often a full answer) that is independent from having completed the questions before in the same exercise.
- In 2024/25, grade 1/6 was at 0/36 points, grade 4/6 was at 14.5/36 points, grade 6/6 at 25/36 points, and the rest interpolated linearly.

.

Belief propagation equations summary

Consider a factor graph representation for the following probability distribution

$$P\left(\{s_i\}_{i=1}^N\right) = \frac{1}{Z}\prod_{i=1}^N g_i(s_i)\prod_{a=1}^M f_a\left(\{s_i\}_{i\in\partial a}\right) .$$

with variable nodes indexed by $i = 1, \dots, N$ and factor nodes indexed by $a = 1, \dots, M$. The Belief Propagation equations read

$$\chi_{s_j}^{j\to a} = \frac{1}{Z^{j\to a}} g_j(s_j) \prod_{b\in\partial j\setminus a} \psi_{s_j}^{b\to j}$$

$$\psi_{s_i}^{a\to i} = \frac{1}{Z^{a\to i}} \sum_{\{s_j\}_{j\in\partial a\setminus i}} f_a\left(\{s_j\}_{j\in\partial a}\right) \prod_{j\in\partial a\setminus i} \chi_{s_j}^{j\to a} .$$

where $Z^{j\to a}$ and $Z^{a\to i}$ are normalization factors set so that $\sum_s \chi_s^{j\to a} = 1$ and $\sum_s \psi_s^{a\to i} = 1$. The Bethe free entropy density Φ is given by

$$N\Phi = \log Z = \sum_{i=1}^N \log Z^i + \sum_{a=1}^M \log Z^a - \sum_{(ia)} \log Z^{ia} ,$$

$$Z^i = \sum_s g_i(s) \prod_{a\in\partial i} \psi_s^{a\to i} ,$$

$$Z^a = \sum_{\{s_i\}_{i\in\partial a}} f_a\left(\{s_i\}_{i\in\partial a}\right) \prod_{i\in\partial a} \chi_{s_i}^{i\to a} ,$$

$$Z^{ia} = \sum_s \chi_s^{i\to a} \psi_s^{a\to i} .$$

The marginal of the variable i is given by

$$\mu_i(s_i) = \frac{1}{Z^i} g_i(s_i) \prod_{a\in\partial i} \psi_{s_i}^{a\to i} .$$

1 Questions (10 points)

1. (1pt) Suppose that $f(x)$ is function with global maximum reached at $x = x_0$. Compute the following limit.

$$\phi = \lim_{N\to\infty} \frac{1}{N} \log \int dx e^{Nf(x)} \,. \tag{1}$$

Solution: By the saddle point approximation, we have

$$\log \int dx e^{Nf(x)} \approx Nf(x_0) \implies \phi = f(x_0)\,.$$

2. (2pt) Consider a linear system of P equations in N variables $\{x_i\}_{i=1}^N$

$$\sum_{i=1}^{N} a_i^\mu x_i = y^\mu \tag{2}$$

for $\mu = 1, \dots, P$, where the coefficients a and y are random. Assume that the matrix of coefficients $a \in \mathbb{R}^{P\times N}$ has maximum rank $\min(N, P)$ for all N. Explain what the SAT/UNSAT transition is and compute it for this problem.

Solution: Thanks to the hypothesis on the rank, the linear system has infinite solutions for $\alpha = P/N < 1$, one solution for $\alpha = P/N = 1$, and zero solution for $\alpha = P/N > 1$ by basic linear algebra counting considerations. Thus the SAT/UNSAT transition, i.e. the value of alpha at which we move from having to not having solutions) is at $\alpha_c = 1$.

3. (2pt) Argue that Bayes optimal inference problems are always Replica Symmetric (RS), i.e.

$$q_{ab} = m_a = m \tag{3}$$

for all replica indices $1 \le a < b \le n$, where q_{ab} is the overlap order parameter and m_a is the magnetization towards the hidden signal.

Solution: By Nishimori identity and magnetization concentration we have that

$$\delta_{m^\star} = \text{Prob}(m_a - m^\star) = \text{Prob}(q_{ab} - m^\star) \tag{4}$$

implying replica symmetry. Here we called $m^\star$ the equilibrium magnetization, and δ_a a Dirac delta at a.

4. (2pt) Explain what a computational hard phase in a Bayes optimal inference problem is. You can suppose that the phase diagram of the problem is given as a function of some signal-to-noise ratio (SNR) parameter, and you can assume that AMP is the best efficient algorithm for the problem at hand. To what kind of phase transition are hard phases related?

Solution: An hard phase is an interval on the SNR axis where the performance of the BO estimator is strictly larger than the performance of AMP.
Hard phases are related to first order phase transitions, as they require at least two local maxima to exist, one dominating the free entropy, and the other describing the performance of AMP.

5. (1pt) Consider the scalar Bayesian inference problem $y \sim P_{\text{out}}(\cdot|x^*)$ where $y \in \mathbb{R}$ is the observation and $x^* \in \mathbb{R}$ is the hidden signal, distributed with prior P_0. Write down the posterior distribution for this problem, and the Bayes optimal estimator w.r.t. the MSE error, both as a function of P_0 and P_{out}.

Solution: The posterior equals

$$P(x|y) = \frac{P_{\text{out}}(y|x)P_0(x)}{\int dx\, P_{\text{out}}(y|x)P_0(x)}$$

and the BO estimator w.r.t. to the MSE is its average, giving

$$\hat{x}(y) = \frac{\int dx\, P_{\text{out}}(y|x)P_0(x)\, x}{\int dx\, P_{\text{out}}(y|x)P_0(x)}$$

6. (2pt) Consider a probability distribution over N variables with pair-wise interactions, i.e.

$$p(x_1, \dots, x_N) = \frac{1}{Z} \prod_{(ij)\in E} f_{(ij)}(x_i, x_j) \tag{5}$$

where E is the edge set of a graph G with N nodes. Argue that if the graph G is a tree (does not have loops), then p admits a representation as a tree factor graph (does not have loops).

Solution: Given that the interactions are pair-wise, one can construct a factor graph representation in which the variable nodes are the nodes of the original graph, and there is exactly one factor node for each of the original edges $(ij) \in E$ placed "in the middle" of the edge, splitting it in two. We then see that the factor graph has the same loop structure as the original graph.

2 Exercise: Computing the largest eigenvalue of a random matrix (22.5 points)

Consider a $N \times N$ symmetric real-valued matrix J. Recall that an eigenvalue of J is a real number λ such that there exists a non-zero vector $v \in \mathbb{R}^N$ satisfying $Jv = \lambda v$. In this exercise, we are interested in computing the value of the largest eigenvalue $\lambda_{\max}$ of the matrix J, when J is a random matrix and $N \to \infty$.

Recall that the largest eigenvalue $\lambda_{\max}$ of a symmetric matrix J satisfies

$$\lambda_{\max} = \frac{1}{r^2} \max_{v \in S_N(r)} v^T J v \,, \tag{6}$$

where $S_N(r)$ is the sphere in $\mathbb{R}^N$ with radius r (the equality above holds for all $r > 0$). We will specify the precise form of J in the following. You can always assume that the matrix J we consider is symmetric and real-valued and has a non-degenerate largest eigenvalue, i.e. there is only one non-null vector $v \in \mathbb{R}^N$ satisfying $Jv = \lambda_{\max} v$. Also, you can assume that $\lambda_{\max}$ converges to a finite positive value for $N \to \infty$.

Finally, always assume that eigenvectors are normalized to have norm $\sqrt{N}$, i.e. to lie on $S_N(\sqrt{N})$.

2.1 General considerations (3.5 points)

7. (0.5pt) Write a function $H_J : S_N(\sqrt{N}) \to \mathbb{R}$ whose global minimum equals $-N\lambda_{\max}$.

Solution:

$$H_J(v) = -v^T J v = -\sum_{i,j=1}^{N} J_{ij} v_i v_j \,,$$

8. (0.5pt) Consider a physical system described by a vector $v \in S_N(\sqrt{N})$ with Hamiltonian given by the function H_J. Write its canonical Gibbs distribution at inverse temperature β.

Solution:

$$p(v) = \frac{e^{-\beta H_J(v)} \delta(v^t v - N)}{Z_J(\beta)}$$

where

$$Z_J(\beta) = \int dv \, \delta(v^t v - N) \, e^{-\beta H_J(v)}$$

9. (0.5pt) What is the ground state of this physical system? What is its energy density (averaged over the associated Gibbs distribution)?

Solution: The ground state of the physical system is the unique (by assumption) eigenvector associated to the largest eigenvalue of J. Its energy density is given by $\lambda_{\max}$.

10. (2 pt) Describe, for fixed J, how to compute $\lambda_{\max}$ in the $N \to \infty$ limit starting from the partition function of the Gibbs distribution defined above. (Just describe the steps and define all important quantities. Do not perform the computation).

Solution: $\lambda_{\max}$ is the ground state energy density of the system. The energy density at finite inverse temperature β can be computed in the $N \to \infty$ limit by computing the free entropy density $\phi(\beta) = N^{-1} \log Z_J$, and taking the derivative

$$e(\beta) = -\partial_\beta \phi(\beta) \, .$$

$\lambda_{\max}$ can then be computed by taking the $\beta \to \infty$ limit.

2.2 Wigner matrices (7pt)

We now specify our random matrix model. We consider a symmetric matrix $J = (G + G^T)/\sqrt{2N}$, where G is a non-symmetric $N \times N$ matrix with independent and identically distributed entries. Each entry is a mean zero, variance one Gaussian random variable. In other words:

- All entries in the lower-triangle of the matrix J_{ij} with $1 \le i < j \le N$ are i.i.d. Gaussians with zero mean and variance $1/N$.
- All entries in the upper-triangle of the matrix J_{ij} with $1 \le j < i \le N$ are determined by symmetry $J_{ij} = J_{ji}$.
- All entries in the diagonal of the matrix J_{ii} with $1 \le i \le N$ are i.i.d. Gaussians with zero mean and variance $2/N$.

We want to compute $\lambda_{\max}$ for a typical matrix J extracted from the random ensemble just defined. To do this, we will start by computing the free entropy density $\phi(\beta) = N^{-1}\mathbb{E}_J \log Z_J(\beta)$ for the specific Hamiltonian

$$H_J(v) = -v^T J v = -\sum_{i,j=1}^{N} J_{ij} v_i v_j \, , \tag{7}$$

where $Z_J(\beta)$ is the partition function of the associated Gibbs distribution, using replica theory. We thus start by computing the integer moments $\mathbb{E}_J Z_J(\beta)^n$, where n is a positive integer.

(Continues on next page)

11. (2pt) Show that the moments $\mathbb{E}_J Z_J(\beta)^n$ can be written as

$$\begin{aligned}\mathbb{E}_J Z_J(\beta)^n = &\int \left(\prod_{1\le a<b\le n} dq_{ab}\right) \exp\left(N\beta^2 \sum_{a,b=1}^{n} q_{ab}^2\right) \\ &\times \int \left(\prod_{a=1}^{n} dv_a\right)\left(\prod_{1\le a<b\le n} \delta\left(v_a^T v_b - N q_{ab}\right)\right)\left(\prod_{a=1}^{n} \delta\left(v_a^T v_a - N\right)\right)\end{aligned} \tag{8}$$

> **Solution:**
>
> $$\begin{aligned}\mathbb{E}_J Z_J(\beta)^n &= \mathbb{E}_J \int \left(\prod_{a=1}^{n} dv_a\right)\left(\prod_{a=1}^{n} \delta\left(v_a^T v_a - N\right)\right) \exp\left(\beta \sum_{i,j=1}^{N} J_{ij} v_i v_j\right) \\ &= \mathbb{E}_G \int \left(\prod_{a=1}^{n} dv_a\right)\left(\prod_{a=1}^{n} \delta\left(v_a^T v_a - N\right)\right) \exp\left(\sqrt{2/N}\beta \sum_{i,j=1}^{N}\sum_{a=1}^{n} G_{ij} v_i^a v_j^a\right) \\ &= \int \left(\prod_{a=1}^{n} dv_a\right)\left(\prod_{a=1}^{n} \delta\left(v_a^T v_a - N\right)\right) \exp\left(\beta^2 N^{-1} \sum_{i,j=1}^{N}\sum_{a,b=1}^{n} v_i^a v_j^a v_i^b v_j^b\right) \\ &= \int \left(\prod_{1\le a<b\le n} dq_{ab}\right) \exp\left(N\beta^2 \sum_{a,b=1}^{n} q_{ab}^2\right) \\ &\quad\times \int \left(\prod_{a=1}^{n} dv_a\right)\left(\prod_{1\le a<b\le n} \delta\left(v_a^T v_b - N q_{ab}\right)\right)\left(\prod_{a=1}^{n} \delta\left(v_a^T v_a - N\right)\right)\end{aligned}$$

12. (0.5pt) What condition does the overlap order parameter $\{q_{ab}\}_{1\le a<b\le n}$ satisfy under the Replica Symmetric (RS) ansatz?

> **Solution:** In the RS ansatz, all out-of-diagonal component of the overlap order parameter are the same, i.e.
>
> $$q_{ab} = q$$
>
> for some real parameter q.

13. (1.5pt) Using the replica trick, show that in the RS ansatz and for $N \to \infty$

$$\phi(\beta) = \text{extr}_q \left(\beta^2(1-q^2) + \frac{1}{2}\log(1-q) + \frac{q}{2(1-q)}\right). \tag{9}$$

You can use (without deriving it) the fact that in the RS ansatz and for large N, small n,

$$\begin{aligned}
&\frac{1}{nN}\log\int\left(\prod_{a=1}^{n}dv_a\right)\left(\prod_{1\le a<b\le n}\delta\left(v_a^T v_b - Nq_{ab}\right)\right)\left(\prod_{a=1}^{n}\delta\left(v_a^T v_a - N\right)\right) \\
&\quad\approx \frac{1}{2}\log(1-q)+\frac{q}{2(1-q)}\,.
\end{aligned} \tag{10}$$

Solution: Under the RS ansatz we have

$$\sum_{a,b=1}^{n} q_{ab}^2 = n + n(n-1)q^2\,.$$

Thus, in the RS ansatz (and using the suggested approximation) and using the saddle point method we get

$$\begin{aligned}
\mathbb{E}_J Z_J(\beta)^n &\approx \int dq\exp\left(nN\left(\beta^2(1-q^2+nq^2)+\frac{1}{2}\log(1-q)+\frac{q}{2(1-q)}\right)\right) \\
&\approx \exp\left(nN\mathrm{extr}_q\left(\beta^2(1-q^2+nq^2)+\frac{1}{2}\log(1-q)+\frac{q}{2(1-q)}\right)\right)\,.
\end{aligned}$$

Using the replica trick (and the fact that $n\to 0$) we finally get

$$\phi(\beta)\approx\frac{\mathbb{E}_J Z_J(\beta)^n-1}{nN}\approx\mathrm{extr}_q\left(\beta^2(1-q^2)+\frac{1}{2}\log(1-q)+\frac{q}{2(1-q)}\right)$$

14. (1pt) Derive and solve the RS state equation for q. If multiple solutions for q arise, consider solutions with $0<q<1$.

Solution: The state equation is given by the condition

$$\partial_q\left(\beta^2(1-q^2)+\frac{1}{2}\log(1-q)+\frac{q}{2(1-q)}\right)=0$$

giving

$$4\beta^2 q=\frac{q}{(1-q)^2}\implies q=1-\frac{1}{2\beta}\,.$$

15. (1pt) Compute the energy density as a function of β (under the RS ansatz).

Solution:
$$e(\beta) = -\partial_\beta \phi(\beta) = -2\beta(1-q^2) = -2 + \frac{1}{2\beta}\,.$$
Here we used that we are at the saddle point to just take the partial derivative of the variational free entropy at fixed order parameter, instead of first substituting and then taking the full derivative.

16. (1pt) Derive the numerical value of the maximum eigenvalue $\lambda_{\max}$ of a typical matrix J from the random ensemble defined above (under the RS ansatz).

Solution: The energy density for $\beta \to \infty$ (we want the largest eigenvalue, i.e. the ground state energy density of the Hamiltonian) equals -2, hence $\lambda_{\max} = 2$.

2.3 Spiked-Wigner matrices (7 points)

Consider now a different random matrix ensemble, namely
$$J = \frac{G + G^T}{\sqrt{2N}} + \frac{\lambda}{N} u u^T \tag{11}$$
where u is a fixed vector on $S_N(\sqrt{N})$, and G is a non-symmetric $N \times N$ matrix with independent and identically distributed entries, where each entry is a mean zero, variance one Gaussian random variable (same as in Section 2.2).

17. (2pt) Show that the moments $\mathbb{E}_J Z_J(\beta)^n$ can be written as

$$\mathbb{E}_J Z_J(\beta)^n = \int \left(\prod_{1\le a<b\le n} dq_{ab}\right) \exp\left(N\beta^2 \sum_{a,b=1}^{n} q_{ab}^2 + N\lambda\beta m_a^2\right) \int \left(\prod_{a=1}^{n} dv_a\right) \times \left(\prod_{1\le a<b\le n} \delta\left(v_a^T v_b - Nq_{ab}\right)\right) \left(\prod_{a=1}^{n} \delta\left(v_a^T v_a - N\right)\right) \left(\prod_{a=1}^{n} \delta\left(v_a^T u - Nm_a\right)\right) \tag{12}$$

Solution:

$$\begin{aligned}
\mathbb{E}_J Z_J(\beta)^n &= \mathbb{E}_J \int \left(\prod_{a=1}^{n} dv_a\right) \left(\prod_{a=1}^{n} \delta\left(v_a^T v_a - N\right)\right) \exp\left(\beta \sum_{i,j=1}^{N} J_{ij} v_i v_j\right) \\
&= \mathbb{E}_G \int \left(\prod_{a=1}^{n} dv_a\right) \left(\prod_{a=1}^{n} \delta\left(v_a^T v_a - N\right)\right) \exp\left(\sqrt{2/N}\beta \sum_{i,j=1}^{N} \sum_{a=1}^{n} G_{ij} v_i^a v_j^a + \beta \frac{\lambda}{N} \sum_{a=1}^{n} \left(\sum_{i=1}^{N} v_i^a u_i\right)^2\right) \\
&= \int \left(\prod_{a=1}^{n} dv_a\right) \left(\prod_{a=1}^{n} \delta\left(v_a^T v_a - N\right)\right) \exp\left(\beta^2 N^{-1} \sum_{i,j=1}^{N} \sum_{a,b=1}^{n} v_i^a v_j^a v_i^b v_j^b + \beta \frac{\lambda}{N} \sum_{a=1}^{n} \left(\sum_{i=1}^{N} v_i^a u_i\right)^2\right) \\
&= \int \left(\prod_{1\le a<b\le n} dq_{ab}\right) \exp\left(N\beta^2 \sum_{a,b=1}^{n} q_{ab}^2 + N\lambda\beta \sum_{a=1}^{n} m_a^2\right) \\
&\quad \times \int \left(\prod_{a=1}^{n} dv_a\right) \left(\prod_{1\le a<b\le n} \delta\left(v_a^T v_b - Nq_{ab}\right)\right) \left(\prod_{a=1}^{n} \delta\left(v_a^T v_a - N\right)\right) \left(\prod_{a=1}^{n} \delta\left(v_a^T u - Nm_a\right)\right)
\end{aligned}$$

Referring back to question 11 to avoid redoing the computation in detail was considered a good answer.

18. (0.5pt) What condition does the overlap order parameter $\{q_{ab}\}_{1\le a<b\le n}$ and the magnetization order parameter $\{m_a\}_{1\le a\le n}$ satisfy under the Replica Symmetric (RS) ansatz?

Solution: In the RS ansatz, all out-of-diagonal component of the overlap order parameter are the same, i.e.

$$q_{ab} = q$$

for some real parameter q. Moreover, all components of the magnetization are the same $m_a = m$, for some real parameter m.

In the RS ansatz, one can show that

$$\phi(\beta) = \mathrm{extr}_{m,q} \left(\beta^2(1-q^2) + \lambda\beta m^2 + \frac{1}{2}\log(1-q) + \frac{q-m^2}{2(1-q)}\right). \tag{13}$$

19. (2pt) Derive and solve the RS state equations for m, q. If multiple solutions for m, q arise, you can assume that the dominant one is the one with largest value of m. In any case, you can discard negative solutions.

Solution: The state equations are given by the conditions

$$\partial_q \left(\beta^2(1-q^2) + \lambda\beta m^2 + \frac{1}{2}\log(1-q) + \frac{q-m^2}{2(1-q)} \right) = 0$$
$$\partial_m \left(\beta^2(1-q^2) + \lambda\beta m^2 + \frac{1}{2}\log(1-q) + \frac{q-m^2}{2(1-q)} \right) = 0$$

giving

$$-2\beta^2 q + \frac{q}{2(1-q)^2} - \frac{m^2}{2(1-q)^2} = 0 \qquad \text{and} \qquad 2\lambda\beta m - \frac{2m}{2(1-q)} = 0$$

The second equations gives either $m = 0$ or

$$q = 1 - \frac{1}{2\lambda\beta}$$

If $m = 0$ the first equation gives

$$q = 1 - \frac{1}{2\beta}$$

If instead $q = 1 - \frac{1}{2\lambda\beta}$ then

$$\begin{aligned} \frac{m^2}{2(1-q)^2} &= -2\beta^2 q + \frac{q}{2(1-q)^2} \\ m^2 &= q\left(1 - 4\beta^2(1-q)^2\right) \\ m^2 &= q\left(1 - 4\beta^2 \frac{1}{4\lambda^2\beta^2}\right) \\ m^2 &= \left(1 - \frac{1}{2\lambda\beta}\right)\left(1 - \frac{1}{\lambda^2}\right) \end{aligned}$$

This solution exists for $\lambda > \max(1, 1/(2\beta))$. Thus,

$$m = \begin{cases} 0 & \text{for } 0 < \lambda < \max(1, 1/(2\beta)) \\ \sqrt{\left(1 - \frac{1}{2\lambda\beta}\right)\left(1 - \frac{1}{\lambda^2}\right)} & \text{for } \lambda > \max(1, 1/(2\beta)) \end{cases}$$
$$q = \begin{cases} 1 - \frac{1}{2\beta} & \text{for } 0 < \lambda < \max(1, 1/(2\beta)) \\ 1 - \frac{1}{2\lambda\beta} & \text{for } \lambda > \max(1, 1/(2\beta)) \end{cases}$$

20. (1pt) Compute the energy density as a function of β (under the RS ansatz).

Solution:

$$\begin{aligned} e(\beta) &= -\partial_\beta \phi(\beta) \\ &= -2\beta(1-q^2) - \lambda m^2 \\ &= -\begin{cases} 2 - \frac{1}{2\beta} & \text{for } 0 < \lambda < \max(1, 1/(2\beta)) \\ \frac{2}{\lambda} - \frac{1}{2\lambda^2\beta} + \lambda\left(1 - \frac{1}{2\lambda\beta}\right)\left(1 - \frac{1}{\lambda^2}\right) & \text{for } \lambda > \max(1, 1/(2\beta)) \end{cases} . \end{aligned}$$

21. (1pt) Derive the numerical value of the maximum eigenvalue $\lambda_{\max}$ of a typical matrix J from the random ensemble defined above as a function of λ (under the RS ansatz).

Solution:

$$\lambda_{\max} = -\lim_{\beta\to\infty} e(\beta) = \begin{cases} 2 & \text{for } 0 < \lambda < 1 \\ \lambda + \frac{1}{\lambda} & \text{for } \lambda > 1 \end{cases} .$$

22. (0.5pt) Does the system undergo a phase transition? If so, of which order?

Solution: Yes, of second order as all order parameters and the energy are continuous.

2.4 A link with inference (5 points)

Consider now the inference problem of retrieving an unknown vector u from the observation

$$J = \frac{G + G^T}{\sqrt{2N}} + \frac{\lambda}{N} u^T u \tag{14}$$

where u is a uniformly sampled vector on $S_N(\sqrt{N})$, and G is a non-symmetric $N \times N$ matrix with independent and identically distributed entries. Each entry is a mean zero, variance one Gaussian random variable.

Recall that the cosine similarity between two vectors on the sphere $S_N(\sqrt{N})$ is given by

$$c(u, v) = \frac{1}{N} u^T v = \frac{1}{N} \sum_{i=1}^{N} v_i u_i \,, \tag{15}$$

and always assume that eigenvectors are normalized to have norm $\sqrt{N}$, i.e. to lie on $S_N(\sqrt{N})$.

23. (1pt) Compute the cosine similarity between the vector u and the eigenvector corresponding to the largest eigenvalue $v_{\max}$ in the limit $N \to \infty$ under the RS ansatz, as a function of λ.

Solution: The cosine similarity is given by $m(\beta)$ of question 19, and that of the maximum eigenvector by the large β limit, giving

$$c(u, v_{\text{max}}) = \begin{cases} 0 & \text{for } 0 < \lambda < 1 \\ \sqrt{1 - \frac{1}{\lambda^2}} & \text{for } \lambda > 1 \end{cases}$$

Partial points were awarded if the link between cosine similarity and the magnetization of question 19 was given, even though the actual equation of the solution was not given.

(Continues on the next page)

24. (2pt) Argue that the Gibbs distribution with Hamiltonian

$$H_J(v) = -v^T J v = -\sum_{i,j=1}^{N} J_{ij} v_i v_j \tag{16}$$

and inverse temperature β is equivalent to the posterior distribution for the Bayes optimal inference problem for the spike u defined above. For which value of β does this equivalence hold?

Solution: The posterior is given by

$$\begin{aligned} P(v|J) &= \frac{1}{Z_J} \prod_{i<j} e^{-\frac{N}{2}(J_{ij} - \frac{\lambda}{N} v_i v_j)^2} \prod_i e^{-\frac{N}{4}(J_{ii} - \frac{\lambda}{N} v_i v_i)^2} \delta(v^T v - N) \\ &= \frac{1}{Z'_J} \prod_{i<j} e^{J_{ij} \lambda v_i v_j} \prod_i e^{-\frac{1}{2} J_{ii} \lambda v_i^2} \delta(v^T v - N) \\ &= \frac{1}{Z'_J} e^{\frac{\lambda}{2} \sum_{ij} J_{ij} v_i v_j} \delta(v^T v - N) \end{aligned}$$

where the quartic term in v is just a function of its norm, hence it is constant. Thus, the Gibbs distribution equals the Bayes optimal posterior for $\beta = \lambda/2$.

25. (2pt) Compute the cosine similarity between the vector u and a sample of the posterior distribution for the Bayes optimal inference problem for the spike u, in the limit $N \to \infty$ under the RS ansatz, as a function of λ.

Solution: The cosine similarity is given by $m(\beta = \lambda/2)$ of question 19. The only non-trivial point to make is that on the computation above we did not average over u, we kept u fixed. If we did average over u, the result would be the same, as the Gibbs distribution is invariant under any rotation of u and v jointly. Back to the computation. First notice that

$$0 < \lambda < \max(1, 1/(2\beta)) \implies 0 < \lambda < \max(1, 1/\lambda)$$

and

$$\lambda < 1/\lambda \iff \lambda^2 < 1 \iff \lambda < 1$$

so that

$$0 < \lambda < \max(1, 1/(2\beta)) \iff 0 < \lambda < 1\,.$$

Thus

$$c(u, v_{\text{Bayes}}) = \begin{cases} 0 & \text{for } 0 < \lambda < 1 \\ 1 - \frac{1}{\lambda^2} & \text{for } \lambda > 1 \end{cases}$$

Again, making the link with question 19 without having explicit solution was awarded partial points.

3 Exercise: The hard-core model (3.5 points)

Given a graph, can we put a large number of particles on the vertices avoiding any first-neighbor contact? This is the task one has to solve in the hard-core model. It is defined as follows. We consider a graph $G = (V, E)$ of size $N = |V|$ and associate an occupation number $\sigma_i \in \{0, 1\}$ to every vertex $i \in V$, where 0 stands for free and 1 for occupied. The Gibbs measure that corresponds to the hard-core model reads

$$P(\{\sigma_i\}_{i=1,\ldots,N}) = \frac{1}{Z} e^{\mu \sum_{i\in V} \sigma_i} \prod_{(ij)\in E} (1 - \sigma_i \sigma_j)\,, \tag{17}$$

where μ is a constant called the chemical potential and Z is the partition function. The parameter μ controls the amount of particles in the graph. As μ increases, the Gibbs measure gives higher probability to configurations with larger number of particles, as measured by the average packing fraction $\rho = 1/N \langle \sum_{j=1}^{N} \sigma_j \rangle$, where angular average are over the distribution (17).

26. (1pt) Provide a factor graph representation of the Gibbs measure in (17).

Solution: The Gibbs measure has only either pairwise interactions or single particle terms. Thus, a valid factor graph representation has the original nodes of the graph as variable nodes, one factor linked to each variable node with weight $e^{\mu\sigma_i}$, and one factor sitting on each edge of the original graph with weight $(1 - \sigma_i \sigma_j)$.

27. (1pt) Write the Belief Propagation (BP) equations for the factor graph you derived above.

Solution: Just specify the BP equations given at the beginning of the exam putting in the fact that factors are on edges (i.e. $a = (ij)$) and that the factor weights are as given in the previous answer.

28. (1.5pt) Suppose that the graph G is tree-like in the $N \to \infty$ limit, and that you found the correct fixed point of the BP equations describing the thermodynamic properties of (17) as a function of μ. Describe how you would compute the average packing fraction using the BP messages.

Solution: We have

$$\rho = \frac{1}{N} \sum_{j=1}^{N} \sum_{\sigma_1=0,1} \cdots \sum_{\sigma_N=0,1} \sigma_j P(\{\sigma_i\}_{i=1,\ldots,N}) = \frac{1}{N} \sum_i \sum_{\tau=0,1} \tau \mu_i(\tau) = \frac{1}{N} \sum_i \mu_i(1)$$

where μ_i is the i-th marginal

$$\mu_i(\tau) = \sum_{\sigma_1=0,1} \cdots \sum_{\sigma_N=0,1} P(\{\sigma_i\}_{i=1,\ldots,N})$$

without summing over $\sigma_i = \tau$. Thus, one can use the expression of the BP marginal to compute all the $\mu_i(1)$, and average them.