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

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

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

$$\begin{aligned}
N\Phi = \log Z &= \sum_{i=1}^{N} \log Z^i + \sum_{a=1}^{M} \log Z^a - \sum_{(ia)} \log Z^{ia} , \\
Z^i &= \sum_s g_i(s) \prod_{a\in\partial i} \psi_s^{a\to i} , \\
Z^a &= \sum_{\{s_i\}_{i\in\partial a}} f_a\left(\{s_i\}_{i\in\partial a}\right) \prod_{i\in\partial a} \chi_{s_i}^{i\to a} , \\
Z^{ia} &= \sum_s \chi_s^{i\to a} \psi_s^{a\to i} .
\end{aligned}$$

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.

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

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.

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

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?

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}.

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

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

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).

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 $-N\lambda_{\max}$.

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.

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

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).

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

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

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?

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

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

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

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

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

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

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).

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

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

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?

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.

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

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

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

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.

(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?

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.

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).

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

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.