UNIVERSITY OF AMSTERDAM

On Statistical Inference and Modelling Data with the Ising Model

Applications to Neuronal Activity and US Supreme Court Voting Patterns

April 28, 2025

Student:
Tycho Stam (13303147)

Lecturer:
dr. C.M.C. de de Mulatier
Course:
Theory of Complex Systems
Course number:
5284THCS6Y

Contents

1 Modeling the activity of a single neuron

Q1. Fig. 1 displays the distribution $P(\tau)$ of time intervals τ between consecutive spikes, as derived from the initial dataset. τ_0 is hereby equal to 1.90ms.

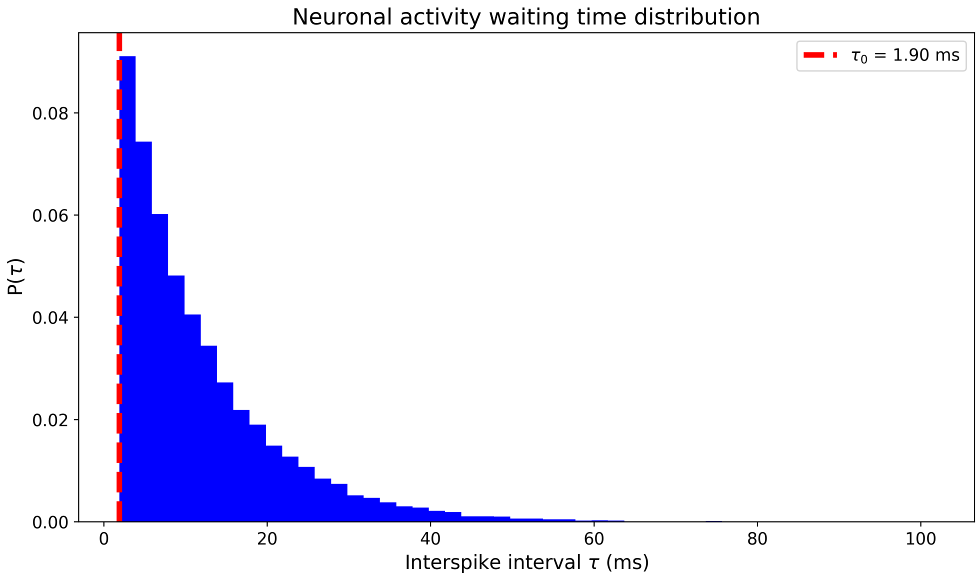

Figure 1: The distribution $P(\tau)$ of time intervals τ between consecutive spikes. Hereby τ_0 shows the refractory period, in which time a neuron does not spike again. For this plot 50 bins were used to group the values.

Q2. When fitting the data on an exponential function, it gives back a pattern similar to the distribution $P(\tau)$, based on $\lambda = 0.08$, as seen in Fig. 2.

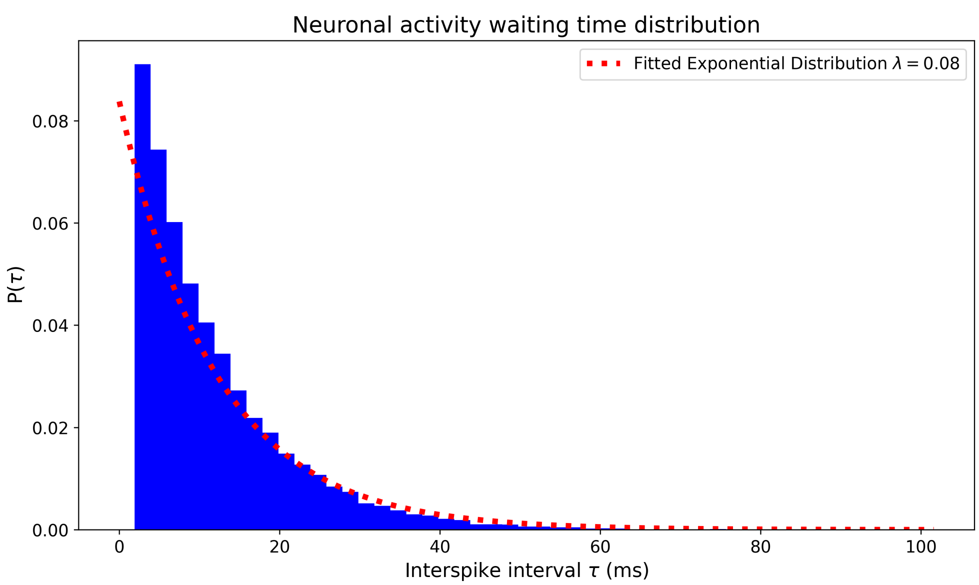

Figure 2: Fitted exponential function of distribution $P(\tau)$, resulting in $\lambda = 0.08$.

Q3. Let τ be the time between two spikes, given what we know about the refractory period we get the following inter-spike interval distribution:

$$P(\tau) = \begin{cases} \lambda \exp\left[-\lambda(\tau - \tau_0)\right] & \text{if } \tau \geq \tau_0 \\ 0 & \text{if } \tau < \tau_0 \end{cases} \tag{1}$$

When comparing the created model with the original data we get the following, as seen in Fig. 3.

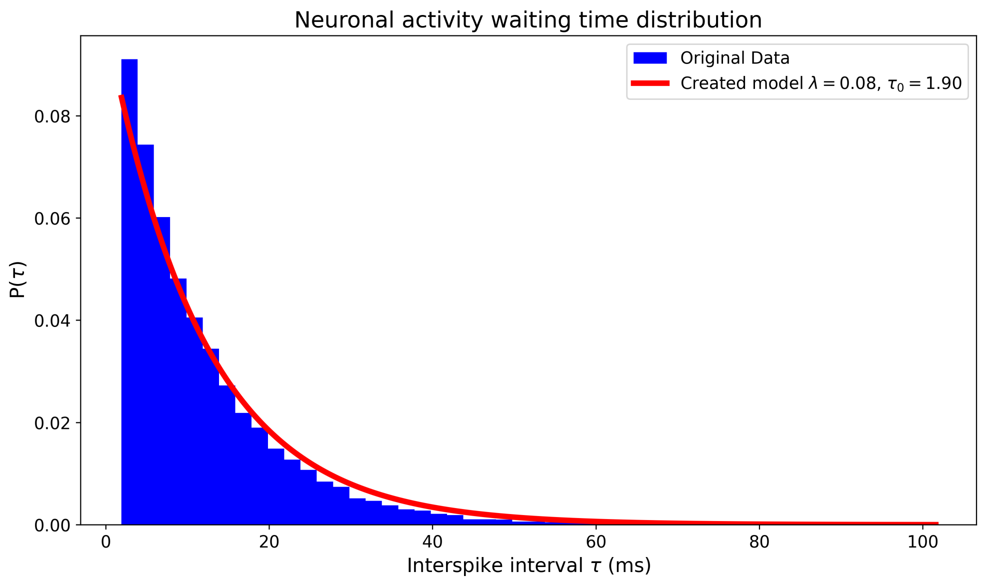

Figure 3: The model data (in red) is plotted over the original dataset, clearly illustrating that the model closely aligns with the observed data.

Q4. To sample from the PDF, we will use inverse transformation sampling. For this the CDF of $P(\tau)$ was created:

$$F(\tau) = \begin{cases} 1 - \exp\left[-\lambda(\tau - \tau_0)\right] & \text{if } \tau \geq \tau_0 \\ 0 & \text{if } \tau < \tau_0 \end{cases} \tag{2}$$

Hereby we sample from $u_k \sim \text{Uniform}(0, 1)$ and solve $u_k = F(\tau)$, giving us:

$$\tau_k = \tau_0 - \frac{\ln(1 - u_k)}{\lambda} \tag{3}$$

Here, k denotes the index of the $k - th$ spike out of a total of $N = 1000$ spikes. This results in Fig. 4.

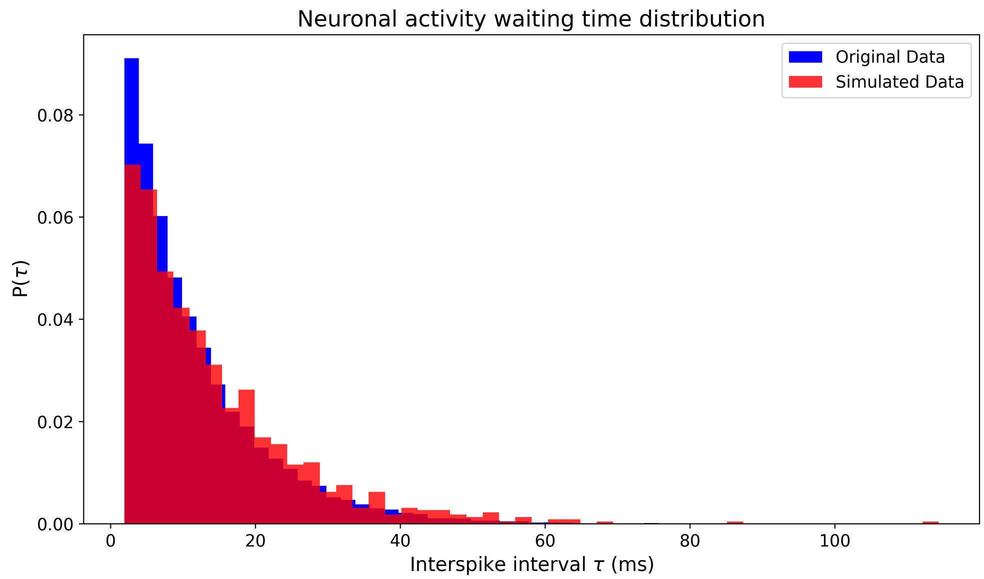

Figure 4: Comparison of the simulated data and original data, sampled using a inverse transformation. Whereby the simulated dataset contains $N = 1000$ samples.

Q5. For the exponential distribution we know that $\mathbb{E}[X] = \frac{1}{\lambda}$. Given Eq. 1, this will result in $\mathbb{E}[\tau] = \frac{1}{\lambda} + \tau_0$. The average firing rate is than defined as:

$$f = \frac{1}{\mathbb{E}[\tau]} = \frac{1}{\tau_0 + \frac{1}{\lambda}} = \frac{\lambda}{1 + \tau_0 \lambda} \tag{4}$$

When plugging in λ and τ_0 we get the average spiking rate of 72.3 Hz

2 Modeling binary data with the Ising model

2.1 Pairwise spin model

For the pairwise spin model we will take the probability distribution to have the general form of an Ising model:

$$p_{\mathbf{g}}(s) = \frac{1}{Z(\mathbf{g})} \exp\left(\sum_{i=1}^{n} h_i s_i + \sum_{pair(i,j)} J_{ij} s_i s_j \right) \tag{5}$$

where n is the number of spin variables, where $pair(i,j)$ denotes a summation over all possible pairs of distinct spin variables, where $g = (h_1, \dots, h_n, J_{1,2}, \dots, J_{n-1,n})$ is a vector of (real) parameters, and where $Z(\mathbf{g})$ is a normalization factor.

Q1.1. Given the sum over $pair(i,j)$, where J_{ij} is the strength of the coupling between. We will have $n-1$ terms, where n is the number of spins in the model. This is because each i will be coupled with all of the other spins, and this scales with the number of spins. Because of the same reasoning the vector $\mathbf{g}$ will have $n + \binom{n}{2}$ parameters. The first term corresponds to the number of h parameters in the model, while $\binom{n}{2}$ represents the number of $J_{i,j}$ interactions.

Q1.2. When $n = 3$ in Eq. 5 we get the following expression:

$$\mathbf{g} = (h_1, h_2, h_3, J_{1,2}, J_{1,3}, J_{2,3})$$

$$p_{\mathbf{g}}(s) = \frac{1}{Z(\mathbf{g})} \exp\left(h_1 s_1 + h_2 s_2 + h_3 s_3 + J_{12} s_1 s_2 + J_{13} s_1 s_3 + J_{23} s_2 s_3\right) \tag{6}$$

Q1.3. The Boltzmann distribution is defined as:

$$P(s) = \frac{\exp[-\beta\, E(s)]}{Z} \tag{7}$$

where $\beta = \frac{1}{k_b T}$. Given that $\beta = 1$, the energy function ($E(s)$) is defined as follows:

$$E(s) = -\left(\sum_{i=1}^{n} h_i s_i + \sum_{pair(i,j)} J_{ij} s_i s_j\right) \tag{8}$$

and the partition function Z:

$$Z = \sum_s \exp\left(\sum_{i=1}^{n} h_i s_i + \sum_{pair(i,j)} J_{ij} s_i s_j\right) \tag{9}$$

The goal of the partition function is to normalize the constants over all s.

Q1.4 h_i represents the external influences on the individual judges, such as policital bias. If h_i is positive, the direction of s_i will also want to be positive, to give the lowest associated energy $-h_i s_i$. Hereby J_{ij} represents the interaction between the connected judges, if it is positive the judges will most likely vote the same way. So if J_{ij} is positive the alignment of the spins should be in the same direction (i.e. both negative or positive), as this will make the product 1 and will minimize the coupling energy $-J_{ij} s_i s_j$.

2.2 Observables

Q2.1. Given the stationary probability distribution of the state $p_{\mathbf{g}}(s)$, we can define $\langle s_i \rangle$ and $\langle s_i s_j \rangle$ as follows:

$$\langle s_i \rangle = \sum_s s_i p_g(s) \tag{10}$$

$$\langle s_i s_j \rangle = \sum_s s_i s_j p_g(s) \tag{11}$$

Q2.2. To calculate the empirical averages denoted by $\langle s_i \rangle_D$ and $\langle s_i s_j \rangle_D$ over the whole dataset, we can use the following equations:

$$\langle s_i \rangle_D = \frac{1}{N} \sum_{k=1}^{N} s_i^k$$

$$\langle s_i s_j \rangle_D = \frac{1}{N} \sum_{k=1}^{N} s_i^k s_j^k$$

Where N is the amount of samples in the dataset $\hat{s}$.

Q2.3. The samples s^k are defined as independent and identically distributed (i.i.d.) and are drawn from the distribution $p(s)$. Given the Law of Large Numbers we can state that:

$$\lim_{N\to\infty} \langle s_i \rangle_D = \mathbb{E}_{p(s)}[s_i] = \langle s_i \rangle \tag{12}$$

The same property holds true for $\langle s_i s_j \rangle_D$, giving us:

$$\lim_{N\to\infty} \langle s_i s_j \rangle_D = \mathbb{E}_{p(s)}[s_i s_j] = \langle s_i s_j \rangle \tag{13}$$

2.3 Maximum Entropy models

Q3.1. The equation for Shannon entropy is defined as follows:

$$S\left[p(s)\right] = -\sum_s p(s)\log p(s) \tag{14}$$

Whereby $p(s) = p_{\mathbf{g}}(s)$ and $p_{\mathbf{g}}(s)$ is defined in Eq. 5.

Q3.2. Based the following sets constraints:

$$\underbrace{\sum_s p(s) = 1}_{\text{normalization constraint}} \quad \text{and} \quad \underbrace{\sum_s p(s)s_i(s) = \langle s_i \rangle_D}_{\text{magnetization constraint}} \quad \text{and} \quad \underbrace{\sum_s p(s)s_i(s)s_j(s) = \langle s_i s_j \rangle_D}_{\text{correlation constraint}} \tag{15}$$

We will define the total number of constraints as m. The normalization constraint contributes 1 constraint. The magnetization constraints contribute n constraints, one for each i. Finally, the correlation constraints contribute $\binom{n}{2}$ constraints, one for each unique pair. Therefore, the total number of constraints is:

$$m = 1 + n + \binom{n}{2}$$

For $n = 3$, this will result in:

$$\begin{aligned} m &= 4 + \frac{3!}{2!(3-2)!} \\ &= 4 + \frac{6}{2} \\ &= 7 \end{aligned}$$

Q3.3. Given the auxiliary function:

$$\begin{aligned} U[p(s)] = S[p(s)] + \lambda_0 \left(\sum_s p(s) - 1\right) + \sum_{i=1}^{n} \alpha_i \left(\sum_s p(s)s_i(s) - \langle s_i \rangle_D\right) \\ + \sum_{pair(i,j)}^{n} \eta_{ij} \left(\sum_s p(s)s_i(s)s_j(s) - \langle s_i s_j \rangle_D\right) \end{aligned} \tag{16}$$

We can substitute Eq. 14 in Eq. 16 to get:

$$\begin{aligned} U[p(s)] = -\sum_s p(s)\log p(s) + \lambda_0 \left(\sum_s p(s) - 1\right) + \sum_{i=1}^{n} \alpha_i \left(\sum_s p(s)s_i(s) - \langle s_i \rangle_D\right) \\ + \sum_{pair(i,j)}^{n} \eta_{ij} \left(\sum_s p(s)s_i(s)s_j(s) - \langle s_i s_j \rangle_D\right) \end{aligned} \tag{17}$$

We will set the fixed state of $p(s)$ to p_s and give it one value of s, removing the summations and giving us:

$$\begin{aligned} U[p_s] = -p_s \log p_s + \lambda_0 \left(p_s - 1\right) + \sum_{i=1}^{n} \alpha_i \left(p_s s_i(s) - \langle s_i \rangle_D\right) \\ + \sum_{pair(i,j)}^{n} \eta_{ij} \left(p_s s_i(s)s_j(s) - \langle s_i s_j \rangle_D\right) \end{aligned} \tag{18}$$

When taking the derivative $\frac{\partial}{\partial p_s}$ over each term in Eq. 18:

$$
\begin{aligned}
\frac{\partial U[p_s]}{\partial p_s} = \frac{\partial}{\partial p_s} - p_s \log p_s + \frac{\partial}{\partial p_s}\lambda_0 \left(p_s - 1\right) + \frac{\partial}{\partial p_s}\sum_{i=1}^{n} \alpha_i \left(p_s s_i(s) - \langle s_i \rangle_D\right) \\
+ \frac{\partial}{\partial p_s} \sum_{pair(i,j)}^{n} \eta_{ij} \left(p_s s_i(s) s_j(s) - \langle s_i s_j \rangle_D\right) \\
= -\log(p_s) - 1 + \lambda_0 + \sum_{i=1}^{n} \alpha_i s_i + \sum_{pair(i,j)}^{n} \eta_{ij} s_i(s) s_j(s)
\end{aligned}
\tag{19}
$$

Proven that the equation stated in the assignment holds.

Q3.4. To maximize the equation we set Eq. 19 to 0:

$$
\begin{aligned}
0 &= \frac{\partial U[p_s]}{\partial p_s} \\
0 &= -\log(p_s) - 1 + \lambda_0 + \sum_{i=1}^{n} \alpha_i s_i + \sum_{pair(i,j)}^{n} \eta_{ij} s_i(s) s_j(s)
\end{aligned}
$$

Solving this equation for p_s gives us:

$$
\begin{aligned}
\log(p_s) &= -1 + \lambda_0 + \sum_{i=1}^{n} \alpha_i s_i + \sum_{pair(i,j)}^{n} \eta_{ij} s_i(s) s_j(s) \\
p_s &= \exp\left(-1 + \lambda_0 + \sum_{i=1}^{n} \alpha_i s_i + \sum_{pair(i,j)}^{n} \eta_{ij} s_i(s) s_j(s)\right) \\
p_s &= \exp\left(-1 + \lambda_0\right) \exp\left(\sum_{i=1}^{n} \alpha_i s_i + \sum_{pair(i,j)}^{n} \eta_{ij} s_i(s) s_j(s)\right)
\end{aligned}
\tag{20}
$$

When substituting the normalization constraint

$$\sum_s p_s = 1$$

in Eq. 20, we get:

$$
\begin{aligned}
\sum_s p_s &= 1 \\
\sum_s \exp\left(-1 + \lambda_0\right) \exp\left(\sum_{i=1}^{n} \alpha_i s_i + \sum_{pair(i,j)}^{n} \eta_{ij} s_i(s) s_j(s)\right) &= 1 \\
\exp\left(-1 + \lambda_0\right) \sum_s \exp\left(\sum_{i=1}^{n} \alpha_i s_i + \sum_{pair(i,j)}^{n} \eta_{ij} s_i(s) s_j(s)\right) &= 1 \\
\exp\left(-1 + \lambda_0\right) &= \frac{1}{\sum_s \exp\left(\sum_{i=1}^{n} \alpha_i s_i + \sum_{pair(i,j)}^{n} \eta_{ij} s_i(s) s_j(s)\right)}
\end{aligned}
$$

Where

$$Z = \sum_s \exp\left(\sum_{i=1}^{n} \alpha_i s_i + \sum_{pair(i,j)}^{n} \eta_{ij} s_i(s) s_j(s)\right)$$

Resulting in:

$$\exp\left(-1+\lambda_0\right)=\frac{1}{Z} \tag{21}$$

and

$$\alpha_i = h_i, \quad \eta_{ij} = J_{ij}$$

2.4 Statistical inference: model with no couplings

Q4.1. We know that the product of powers with the same base is equal to the base raised to the sum of the exponents, based on this relation we can rewrite

$$p_{\mathbf{g}}(s) = \frac{1}{Z(\mathbf{g})} \exp\left(\sum_{i=1}^{n} h_i s_i\right)$$

as:

$$p_{\mathbf{g}}(s) = \frac{1}{Z(\mathbf{g})} \prod_{i=1}^{n} \exp\left(h_i s_i\right)$$

We can rewrite that as:

$$p_{\mathbf{g}}(s) = \prod_{i=1}^{n} p_{h_i}(s_i) \tag{22}$$

Given that the distribution for each spin is normalized, a corresponding partition function Z_i exists for each spin:

$$Z_i = \exp(h_i) + \exp(-h_i) = 2\cosh(h_i)$$

Results in:

$$p_{h_i} \equiv \frac{\exp(h_i s_i)}{2\cosh(h_i)} \tag{23}$$

Q4.2. As stated the probability distribution p_{h_i} is the distribution for a spin variable s_i, resulting in:

$$\langle s_i \rangle = \sum_{s} p_{h_i} s_i$$

Where $s \in \{-1, 1\}$. Using Eq. 23 and the properties of the spin, we will get:

$$\begin{aligned}
\langle s_i \rangle &= (1)p_{h_i} + (-1)p_{h_i} \\
&= \frac{\exp(h_i)}{2\cosh(h_i)} - \frac{\exp(-h_i)}{2\cosh(h_i)} \\
&= \frac{\exp(h_i)}{\exp(h_i)+\exp(-h_i)} - \frac{\exp(-h_i)}{\exp(h_i)+\exp(-h_i)} \\
&= \frac{\exp(h_i)-\exp(-h_i)}{\exp(h_i)+\exp(-h_i)} \\
&= \tanh(h_i)
\end{aligned} \tag{24}$$

Applying the inverse of the function in Eq. 24 yields:

$$h_i = \tanh^{-1}\left(\langle s_i \rangle\right)$$

Given the constraint $\langle s_i \rangle = \langle s_i \rangle_D$

$$h_i = \tanh^{-1}\left(\langle s_i \rangle_D\right) \tag{25}$$

Q4.3. Given that when $\langle s_i \rangle_D > 0$, h_i will be positive, indicating that the i-th judge tends to vote positively overall, a similar pattern as described in question 1.4 will emerge. However, now each judge's vote will be influenced by the average votes of all the other judges.

2.5 Statistical inference: maximizing the log-likelihood function

Q5.1. The dataset $\hat{s} = (s^1, s^2, \ldots, s^N$ consists of N independent samples drawn from the distribution $p_{\mathbf{g}}(s)$. Because of this property, we can rewrite the probability that the model $p_{\mathbf{g}}(s)$ produces the dataset

$$\mathcal{L} = \log P_s(\hat{s})$$

as:

$$\begin{aligned}\mathcal{L} &= \log \left[\prod_{k=1}^{N} p_{\mathbf{g}} \left(s^k \right) \right] \\ &= \sum_{k=1}^{N} \log \left[p_{\mathbf{g}} \left(s^k \right) \right]\end{aligned}$$

Instead of summing over all N, we can sum over the unique states s, resulting in a similar sum:

$$\mathcal{L} = \sum_{s} K(s) \log \left[p_{\mathbf{g}} \left(s^k \right) \right] \tag{26}$$

By definition we know that:

$$\begin{aligned} p_D &= \frac{K(s)}{N} \\ K(s) &= N p_D \end{aligned}$$

When substituting this in Eq. 26 we get:

$$\begin{aligned}\mathcal{L} &= \sum_{s} N P_D \log \left[p_{\mathbf{g}} \left(s^k \right) \right] \\ &= N \sum_{s} P_D \log \left[p_{\mathbf{g}} \left(s^k \right) \right]\end{aligned} \tag{27}$$

Q5.2. First we will rewrite the log-likelihood function in terms of the Ising model:

$$\begin{aligned}\mathcal{L} &= N \sum_{s} p_D \log p_{\mathbf{g}}(s) \\ &= N \sum_{s} p_D \left(\sum_{i=1}^{n} h_i s_i + \sum_{pair(i,j)}^{n} J_{ij} s_i s_j - \log(Z) \right)\end{aligned} \tag{28}$$

First we will take the derivative of Eq. 28 with a respect to h_i:

$$\frac{\partial \mathcal{L}}{\partial h_i} = N \sum_{s} p_D \frac{\partial}{\partial h_i} \left(\sum_{i=1}^{n} h_i s_i + \sum_{pair(i,j)}^{n} J_{ij} s_i s_j - \log(Z) \right) \tag{29}$$

Hereby the first term:

$$\frac{\partial}{\partial h_i} \sum_{i=1}^{n} h_i s_i = s_i$$

The second term:

$$\frac{\partial}{\partial h_i} \sum_{pair(i,j)}^{n} J_{ij} s_i s_j = 0$$

And the last term, given Eq. 9:

$$\begin{aligned}\frac{\partial}{\partial h_i} - \log(Z) &= \frac{1}{Z}\frac{\partial Z}{\partial h_i} \\ &= \frac{1}{Z}\frac{\partial}{\partial h_i}\left[\sum_s \exp\left(\sum_{i=1}^{n} h_i s_i + \sum_{pair(i,j)}^{n} J_{ij} s_i s_j\right)\right] \\ &= \frac{1}{Z}\left[\sum_s s_i \exp\left(\sum_{i=1}^{n} h_i s_i + \sum_{pair(i,j)}^{n} J_{ij} s_i s_j\right)\right] \\ &= \sum_s s_i p_{\mathbf{g}}(s) \\ &= \langle s_i \rangle\end{aligned} \tag{30}$$

Making that:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial h_i} &= N \sum_s p_D(s_i - \langle s_i \rangle \\ &= N\left(\langle s_i \rangle_D - \langle s_i \rangle\right)\end{aligned} \tag{31}$$

When setting the maximum of $\mathcal{L}$:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial h_i} &= 0 \\ N\left(\langle s_i \rangle_D - \langle s_i \rangle\right) &= 0 \\ \left(\langle s_i \rangle_D - \langle s_i \rangle\right) &= 0 \\ \langle s_i \rangle_D &= \langle s_i \rangle\end{aligned} \tag{32}$$

The same can be done with taking the derivative of Eq. 28 with a respect to J_{ij}, giving us the first term:

$$\frac{\partial}{\partial J_{ij}} \sum_{i=1}^{n} h_i s_i = 0$$

The second term:

$$\frac{\partial}{\partial J_{ij}} \sum_{pair(i,j)}^{n} J_{ij} s_i s_j = s_i s_j$$

and

$$\frac{\partial}{\partial h_i} - \log(Z) = \langle s_i s_j \rangle \tag{33}$$

for the third term with the same steps as Eq. 30, but than with a respect to J_{ij}. Giving us the maximum:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial h_i} &= 0 \\ N\left(\langle s_i s_j \rangle_D - \langle s_i s_j \rangle\right) &= 0 \\ \left(\langle s_i s_j \rangle_D - \langle s_i s_j \rangle\right) &= 0 \\ \langle s_i s_j \rangle_D &= \langle s_i s_j \rangle\end{aligned} \tag{34}$$

3 Application to the analysis of the US supreme Court

Q6.1.

Q6.2.

Q6.3.

Q6.4.

Q6.5.

Q6.6.

Q6.7.

Q6.8.

Q6.9.