Lecture 8: Replica method, I: the spiked GOE matrix

Lecturer: Song Mei *Scriber: Taejoo Ahn*

In this lecture, we finish the field theoretic calculation to derive the large deviation of overlap matrix from last lecture, and introduce the replica method using the example of the spiked GOE matrix model. Throughout this lecture, we rely on field theoretic calculation (which is not mathematically rigorous), and thus abuse the use of '=' multiple times, without justification in changing the order of limits, expectation and integrals. We will further justify our main results rigorously on later lectures.

1 Large deviation of overlap matrix and field theoretic calculations.

Remind the setup from the previous lecture. We have $\boldsymbol{\sigma}_1, \ldots, \boldsymbol{\sigma}_k \stackrel{i.i.d.}{\sim} Unif(\mathbb{S}^{n-1}(\sqrt{n}))$, and

$$\boldsymbol{\sigma} := [\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, \ldots, \boldsymbol{\sigma}_k] \in \mathbb{R}^{n\times k}$$

and

$$\bar{\boldsymbol{Q}}(\boldsymbol{\sigma}) := \boldsymbol{\sigma}^\top \boldsymbol{\sigma}/n = \begin{bmatrix} \|\boldsymbol{\sigma}_1\|_2^2/n & \cdots & \langle \boldsymbol{\sigma}_1, \boldsymbol{\sigma}_k\rangle/n \\ \vdots & \ddots & \vdots \\ \langle \boldsymbol{\sigma}_k, \boldsymbol{\sigma}_1\rangle/n & \cdots & \|\boldsymbol{\sigma}_k\|_2^2/n \end{bmatrix} \in \mathbb{R}^{k\times k}$$

Note that $\bar{\boldsymbol{Q}}(\boldsymbol{\sigma})$ is symmetric and that $\bar{\boldsymbol{Q}}(\boldsymbol{\sigma})_{ii} = 1$ for $1 \le i \le k$. Now let $\boldsymbol{Q} \in \mathbb{R}^{k\times k}$ be a symmetric matrix with $\boldsymbol{Q}_{ii} = 1$. Here we are interested in the large deviation of $\bar{\boldsymbol{Q}}$ matrix:

$$\lim_{\epsilon\to 0} \lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(\bar{Q}(\boldsymbol{\sigma})_{ij} \in [-\epsilon + Q_{ij}, Q_{ij} + \epsilon], \forall i, j)$$

Here we can express the 'probability of $\bar{\boldsymbol{Q}}$ being close to $\boldsymbol{Q}$' as below

$$\mathbb{P}(\bar{\boldsymbol{Q}}(\boldsymbol{\sigma}) \approx \boldsymbol{Q}) \doteq \frac{\int_{\mathbb{R}^{n\times k}} \delta(\bar{\boldsymbol{Q}}(\boldsymbol{\sigma}) - \boldsymbol{Q}) \prod_{i=1}^k d\boldsymbol{\sigma}_i}{\int_{\mathbb{R}^{n\times k}} \prod_{i=1}^k \delta(\|\boldsymbol{\sigma}_i\|_2^2 - 1) \prod_{i=1}^k d\boldsymbol{\sigma}_i} \doteq \frac{S_n(\boldsymbol{Q})}{T_n}$$

where the numerator can represent the probability (if $d\boldsymbol{\sigma}_i$ are Lebesgue measures), and the denominator can be regarded as a normalizing constant. Now we will calculate each $S_n(\boldsymbol{Q})$ and T_n. Note that our interest lies on calculating the limit of $\frac{1}{n}\log S_n(\boldsymbol{Q})$ where k is fixed and $n \to \infty$.

$$
\begin{aligned}
S_n(\boldsymbol{Q}) = &\int_{\mathbb{R}^{n\times k}} \prod_{1\le i\le j\le k} \delta(\langle \boldsymbol{\sigma}_i, \boldsymbol{\sigma}_j\rangle - nQ_{ij}) \prod_{i=1}^{k} d\boldsymbol{\sigma}_i \\
&\overset{(1)}{=} \int_{\mathbb{R}^{n\times k}} \prod_{i=1}^{k} d\boldsymbol{\sigma}_i \frac{1}{(2\pi)^{k(k+1)/2}} \int_{\mathbb{R}^{k(k+1)/2}} [\prod_{1\le i\le j\le k} \exp\{-i\lambda_{ij}\langle \boldsymbol{\sigma}_i, \boldsymbol{\sigma}_j\rangle + i\lambda_{ij} nQ_{ij}\} \prod_{1\le i\le j\le k} d\lambda_{ij} \\
&\quad \cdot \prod_{1\le i\le k} \exp\{-i\lambda_{ii}\|\boldsymbol{\sigma}_i\|_2^2/2 + i\lambda_{ii} nQ_{ii}/2\} \prod_{1\le i\le k} d\lambda_{ii}] \\
&\overset{(2)}{\doteq} \inf_{\Lambda\in\mathbb{R}^{k\times k}} \int_{\mathbb{R}^{n\times k}} \left(\prod_{i=1}^{k} d\boldsymbol{\sigma}_i\right) \exp\{-\sum_{i,j=1}^{k} \lambda_{ij}\langle \boldsymbol{\sigma}_i, \boldsymbol{\sigma}_j\rangle/2 + n\sum_{i,j=1}^{k} \lambda_{ij} Q_{ij}/2\} \\
&\overset{(3)}{=} \inf_{\Lambda\in\mathbb{R}^{k\times k}} \int_{\mathbb{R}^{n\times k}} \prod_{i=1}^{k}\prod_{\alpha=1}^{n} d\sigma_i^{\alpha} \exp\{-\sum_{i,j=1}^{k} \lambda_{ij} \sum_{\alpha=1}^{n} \sigma_i^{\alpha}\sigma_j^{\alpha}/2\} \times \exp\{n\sum_{i,j=1}^{k} \lambda_{ij} Q_{ij}/2\} \\
&\overset{(4)}{=} \inf_{\Lambda\in\mathbb{R}^{k\times k}} \left(\int_{\mathbb{R}^{k}} \prod_{i=1}^{k} d\sigma_i \exp\{-\sum_{i,j=1}^{k} \lambda_{ij}\sigma_i\sigma_j/2\}\right)^{n} \times \exp\{n\sum_{i,j=1}^{k} \lambda_{ij} Q_{ij}/2\} \\
&\overset{(5)}{=} \inf_{\Lambda\in\mathbb{R}^{k\times k}} \left(\det(\Lambda)^{-\frac{1}{2}} \cdot (\sqrt{2\pi})^{k}\right)^{n} \times \exp\{n\langle \Lambda, \boldsymbol{Q}\rangle/2\} \\
&= \inf_{\Lambda\in\mathbb{R}^{k\times k}} \exp\{n[\langle \Lambda, \boldsymbol{Q}\rangle/2 - \frac{1}{2}\log\det(\Lambda) + \frac{k}{2}\log 2\pi]\}
\end{aligned}
$$

Here (1) is from the delta identity formula, and (2) is from the saddle point approximation since integration respect to Λ was low dimensional integration. Also note that we dropped the constant factor as it vanishes when we calculate $\frac{1}{n}\log S_n(\boldsymbol{Q})$ and let $n \to \infty$. (3) and (4) are just from rewriting inner product to element-wise expression and factorizing it, and (5) is from the Gaussian formula:

$$
\int_{\mathbb{R}^{k}} \prod_{i=1}^{k} d\sigma_i \exp\{-\sum_{i,j=1}^{k} \lambda_{ij}\sigma_i\sigma_j/2\} = \int_{\mathbb{R}^{k}} \exp\{-\langle \bar{\boldsymbol{\sigma}}, \Lambda\bar{\boldsymbol{\sigma}}\rangle/2\} d\bar{\boldsymbol{\sigma}} = \det(\Lambda)^{-\frac{1}{2}}(\sqrt{2\pi})^{k}
$$

Therefore with simple matrix calculus, we conclude:

$$
\frac{1}{n}\log S_n(\boldsymbol{Q}) = \inf_{\Lambda}[\langle \Lambda, \boldsymbol{Q}\rangle/2 - \frac{1}{2}\log\det(\Lambda) + \frac{k}{2}\log 2\pi] = \frac{1}{2}\log\det(\boldsymbol{Q}) + \frac{k}{2}\log 2\pi
$$

Finally, using the fact that when $Q \approx \mathbb{E}[\bar{\boldsymbol{Q}}(\boldsymbol{\sigma})] = I_k$, the probability should approximately be 1, or

$$
\sup_{Q} \lim_{n\to\infty} \frac{1}{n}\log \mathbb{P}(\bar{\boldsymbol{Q}}(\boldsymbol{\sigma}) \approx \boldsymbol{Q}) = 0
$$

so we can conclude that $\frac{1}{n}\log T_n = \frac{k}{2}\log 2\pi$, and have:

$$
\frac{1}{n}\log \mathbb{P}(\bar{\boldsymbol{Q}}(\boldsymbol{\sigma}) \approx \boldsymbol{Q}) = \frac{1}{2}\log\det(\boldsymbol{Q}) \tag{1}
$$

2 Field theoretic calculation for general large deviations.

In this section we present a general recipe for deriving large deviations using field theoretic calculation. Here we have $\boldsymbol{x}_i \in \mathbb{R}^k$ with $\boldsymbol{x}_i \overset{i.i.d}{\sim} \mathbb{P}_{\boldsymbol{x}}$ and a function $M : \mathbb{R}^k \to \mathbb{R}^p$, and our interest is to calculate the large deviation:

$$\lim_{\epsilon\to 0}\lim_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\|\frac{1}{n}\sum_{i=1}^{n}M(\boldsymbol{x}_i)-m\|_2^2\leq\epsilon\right)$$

We start with

$$\begin{aligned}
\mathbb{P}(\frac{1}{n}\sum_{i=1}^{n}M(\boldsymbol{x}_i)\approx m) &\doteq \mathbb{E}_{\boldsymbol{x}}[\delta(\sum_{i=1}^{n}M(\boldsymbol{x}_i)-nm)]\\
&\overset{(6)}{\doteq}\mathbb{E}_{\boldsymbol{x}}\left[\frac{1}{(2\pi)^p}\int_{\mathbb{R}^p}\exp\{i\langle\sum_{i=1}^{n}M(\boldsymbol{x}_i)-nm,\Lambda\rangle\}d\Lambda\right]\\
&\overset{(7)}{=}\frac{1}{(2\pi)^p}\int_{\mathbb{R}^p}\exp\{-in\langle m,\Lambda\rangle\}\mathbb{E}_{\boldsymbol{x}}\left[\exp\{i\langle\sum_{i=1}^{n}M(\boldsymbol{x}_i),\Lambda\rangle\}\right]d\Lambda\\
&\overset{(8)}{=}\int_{\mathbb{R}^p}\exp\{-in\langle m,\Lambda\rangle\}\left(\mathbb{E}\left[e^{i\langle\Lambda,M(\boldsymbol{x})\rangle}\right]\right)^n d\Lambda\\
&\overset{(9)}{=}\mathrm{ext}_\Lambda\left\{\exp\{-n\langle\Lambda,m\rangle\}\times\left(\mathbb{E}\left[e^{\langle\Lambda,M(\boldsymbol{x})\rangle}\right]\right)^n\right\}
\end{aligned}$$

where (6) is from the delta identity formula and (7) is abuse of changing order of integration and expectation. (8), we dropped the constant factor (with same reason from overlap matrix example) and factorized the integration. (9) is from the saddle point approximation again, note that we replaced $i\Lambda$ with Λ since the extreme value are chosen in complex values of Λ. Finally, taking logarithm and limit, we get

$$\frac{1}{n}\log\mathbb{P}(\frac{1}{n}\sum_{i=1}^{n}M(\boldsymbol{x}_i)\approx m)\doteq\mathrm{ext}_\Lambda\left\{-\langle\Lambda,m\rangle+\log\mathbb{E}\left[e^{\langle\Lambda,M(\boldsymbol{x})\rangle}\right]\right\} \tag{2}$$

3 The Cramer's theorem.

Here we present a theorem that gives the rigorousness of the previous section 's result.

Theorem 1 **(Cramer).** *Let* $X_i\overset{i.i.d}{\sim}\mu_X$*, and* $f:\mathcal{X}\to\mathbb{R}$*. If* $A\subseteq\mathbb{R}$ *is a closed interval of real axis, then*

$$\lim_{n\to\infty}\log\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n}f(X_i)\in A\right)=-\inf_{a\in A}I(a)$$

where

$$I(a):=\sup_\lambda\left\{\lambda a-\log\mathbb{E}_{X\sim\mu_X}\left[e^{\lambda f(X)}\right]\right\}.$$

We also write

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n}f(X_i)\in A\right)\doteq\exp\{-n\inf_{a\in A}I(a)\}$$

We call $I(a)$ the **rate function**, which is the dual of the log-moment generating function of $f(X)$. We will not give the full proof of the theorem in this lecture. Instead, we suggest an intuition which comes from straightforward calculation using the Markov's inequality: For $A=[a,\infty]$, we have

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n}f(X_i)\in A\right)=\mathbb{P}\left(e^{\lambda\sum_{i=1}^{n}f(X_i)}\geq e^{\lambda na}\right)\leq\inf_{\lambda\geq 0}\frac{\mathbb{E}_X\left[e^{\lambda f(X)}\right]^n}{e^{n\lambda a}}$$

4 The replica method.

In this section, we present the replica method, which is a powerful tool to calculate various objects in statistical physics. Here we start with an example of the spiked GOE matrix to illustrate the replica method. Recall the setup of the spiked GOE matrix model: $\boldsymbol{u} \sim Unif(\mathbb{S}^{n-1})$ and $\boldsymbol{Y} := \lambda \boldsymbol{u}\boldsymbol{u}^\top + \boldsymbol{W} \in \mathbb{R}^{n\times n}$ where the signal to noise ratio $\lambda \geq 0$ and the noise matrix $\boldsymbol{W} \sim \mathrm{GOE}(n)$. Our interest lies on calculating the limiting free entropy density

$$\phi(\lambda) := \lim_{n\to\infty} \mathbb{E}\left[\sup_{\boldsymbol{\sigma}\in\mathbb{S}^{n-1}} \langle \boldsymbol{\sigma}, \boldsymbol{Y}\boldsymbol{\sigma}\rangle\right]$$

and the overlap of MLE with the signal vector

$$m(\lambda) := \lim_{n\to\infty} \mathbb{E}\left[\langle \hat{\boldsymbol{\theta}}, \boldsymbol{u}\rangle\right]$$

where $\hat{\boldsymbol{\theta}} := \arg\max_{\boldsymbol{\sigma}\in\mathbb{S}^{n-1}} \langle \boldsymbol{\sigma}, \boldsymbol{Y}\boldsymbol{\sigma}\rangle$ is the MLE to estimate $\boldsymbol{u}$. We have well known closed form of both of them, which we call the **BBP phase transition** formula:

$$\phi(\lambda) = \begin{cases} 2, & \lambda \leq 1, \\ \lambda + \frac{1}{\lambda}, & \lambda > 1, \end{cases}$$

$$m(\lambda) = \begin{cases} 0, & \lambda \leq 1, \\ 1 - \frac{1}{\lambda^2}, & \lambda > 1. \end{cases}$$

Our goal here is to derive the formula using replica method and field theoretic calculations. Recall the free energy approach from lecture 5. Our idea was to find a perturbed Hamiltonian so that we can express $\phi(\lambda)$ as a low temperature limit of an ensemble average of some observable. We have $\Omega = \mathbb{S}^{n-1}$ as the configuration space, and the reference measure ν_0 is uniform on Ω. Then we defined Hamiltonian and potentials as below

$$\begin{aligned} H_\lambda(\boldsymbol{\sigma}) &= -n\langle \boldsymbol{\sigma}, \boldsymbol{W}\boldsymbol{\sigma}\rangle - n\lambda\langle \boldsymbol{\sigma}, \boldsymbol{u}\rangle^2 \\ Z_n(\beta,\lambda) &= \int_\Omega \exp\{-\beta H_\lambda(\boldsymbol{\sigma})\}\nu_0(d\boldsymbol{\sigma}) \\ \Phi_n(\beta,\lambda) &= \log Z_n(\beta,\lambda) \\ \phi(\beta,\lambda) &= \lim_{n\to\infty} \frac{1}{n}\mathbb{E}[\log Z_n(\beta,\lambda)] \end{aligned}$$

Then taking the low temperature limit of the free entropy density gives

$$\begin{cases} \phi(\lambda) = \mathbb{E}[\max_{\boldsymbol{\sigma}} H_\lambda(\boldsymbol{\sigma})/n] = \lim_{\beta\to\infty} \frac{1}{\beta}\phi(\beta,\lambda) \\ m(\lambda) = \phi'(\lambda) \end{cases}$$

The crucial part of this calculation is calculating the free entropy density $\phi(\beta,\lambda)$. The main difficulty comes from the nonlinearity of logarithm, so that we cannot insert the expectation inside the integration by changing the order. Here we use the **replica trick** with heuristic derivation to resolve this issue.

Lemma 2 (Replica trick). $\mathbb{E}[\log Z] \equiv \lim_{k\to 0} \frac{1}{k}\mathbb{E}[Z^k]$.

Heuristic derivation : Note that $\log(1+x) \approx x$ as $x \to 0$. We have

$$\begin{aligned} \mathbb{E}[\log Z] &= \mathbb{E}[(\log Z^k)/k] = \lim_{k\to 0} \mathbb{E}[\log(1 + (Z^k - 1))/k] \\ &= \lim_{k\to 0} \mathbb{E}[(Z^k-1)/k] = \lim_{k\to 0} (\mathbb{E}[Z^k] - 1)/k \\ &= \lim_{k\to 0} \frac{1}{k}\log(1 + \mathbb{E}[Z^k] - 1) = \lim_{k\to 0} \frac{1}{k}\log \mathbb{E}[Z^k] \end{aligned}$$

Now return to the spiked GOE matrix model. Using replica trick, we start from

$$\lim_{n\to\infty}\frac{1}{n}\mathbb{E}[\log Z_n]=\lim_{n\to\infty}\lim_{k\to 0}\frac{1}{nk}\mathbb{E}[Z_n^k]=\lim_{k\to 0}\lim_{n\to\infty}\frac{1}{nk}\mathbb{E}[Z_n^k].$$

To calculate $\mathbb{E}[Z_n^k]$, we use one more trick here: assume k as a positive integer and get closed form of $\mathbb{E}[Z_n^k]$ as a function of k, and then take the limit $k\to 0$ as if the formula works for all real numbers k. Overall steps of the calculation are as below:

1. $S(k,\beta,\lambda):=\lim_{n\to\infty}\frac{1}{n}\log\mathbb{E}[Z_n^k]$
2. $\phi(\beta,\lambda)=\lim_{k\to 0}\frac{1}{k}S(k,\beta,\lambda)$
3. $\phi(\lambda)=\lim_{\beta\to\infty}\frac{1}{\beta}\phi(\beta,\lambda)$

The following lemma helps calculating of the first step.

Lemma 3. *For $k\in\mathbb{N}_+$,*

$$S(k,\beta,\lambda)=\sup_{\boldsymbol{Q}\in\mathbb{R}^{(k+1)\times(k+1)}} U(\boldsymbol{Q})$$
$$s.t.\ \boldsymbol{Q}\succeq 0, Q_{ii}=1$$

where

$$U(\boldsymbol{Q})=\beta\lambda\sum_{i=1}^{k}q_{0i}^2+\beta^2\sum_{i,j=1}^{k}q_{ij}^2+\frac{1}{2}\log\det\boldsymbol{Q}$$

with $\boldsymbol{Q}=(q_{ij})_{0\le i,j\le k}$

Sketch of proof : we start from

$$\begin{aligned}
\mathbb{E}[Z_n(\beta,\lambda)^k]&=\mathbb{E}\left[\left(\int_{\mathbb{S}^{n-1}}\exp\{-\beta H_\lambda(\boldsymbol{\sigma})\}\nu_0(d\boldsymbol{\sigma})\right)^k\right]\\
&=\mathbb{E}\left[\int_{(\mathbb{S}^{n-1})^{\otimes k}}\exp\{-\beta\sum_{a=1}^{k}H_\lambda(\boldsymbol{\sigma}^a)\}\prod_{a=1}^{k}\nu_0(d\boldsymbol{\sigma}^a)\right]\\
&=\int_{(\mathbb{S}^{n-1})^{\otimes k}}\mathbb{E}\left[\exp\{-\beta\sum_{a=1}^{k}H_\lambda(\boldsymbol{\sigma}^a)\}\right]\prod_{a=1}^{k}\nu_0(d\boldsymbol{\sigma}^a)\\
&=\int_{(\mathbb{S}^{n-1})^{\otimes k}}\mathbb{E}\left[\exp\{\beta n(\sum_{a=1}^{k}\lambda\langle\boldsymbol{\sigma}^a,\boldsymbol{u}\rangle^2+\langle\boldsymbol{\sigma}^a,\boldsymbol{W}\boldsymbol{\sigma}^a\rangle)\}\right]\prod_{a=1}^{k}\nu_0(d\boldsymbol{\sigma}^a)\\
&=\int_{(\mathbb{S}^{n-1})^{\otimes k}}\exp\{\beta n(\sum_{a=1}^{k}\lambda\langle\boldsymbol{\sigma},\boldsymbol{u}\rangle^2)\}\cdot\underbrace{\mathbb{E}[\exp\{\beta n\sum_{a=1}^{k}\langle\boldsymbol{\sigma}^a,\boldsymbol{W}\boldsymbol{\sigma}^a\rangle\}]}_{=:\boldsymbol{E}}\prod_{a=1}^{k}\nu_0(d\boldsymbol{\sigma}^a)
\end{aligned}$$

Here we can calculate the expectation part, which we denoted $\boldsymbol{E}$, by using moment generating function of Gaussian. Define a random matrix $\boldsymbol{G}\in\mathbb{R}^{n\times n}$ with each element are independent standard Gaussian random variables. Then $\boldsymbol{W} \stackrel{d}{\equiv} (\boldsymbol{G}+\boldsymbol{G}^{\mathsf{T}})/\sqrt{2n}$, so that we have

$$\begin{aligned}
\boldsymbol{E} &= \mathbb{E}\left[\exp\{\beta n \sum_{a=1}^{k} \langle \boldsymbol{\sigma}^a, (\boldsymbol{G}+\boldsymbol{G}^{\mathsf{T}})\boldsymbol{\sigma}^a\rangle/\sqrt{2n})\}\right] \\
&= \mathbb{E}\left[\exp\left\{\beta n \cdot \mathrm{tr}\left(\sum_{a=1}^{k} \boldsymbol{\sigma}^a(\boldsymbol{\sigma}^a)^{\mathsf{T}}(\boldsymbol{G}+\boldsymbol{G}^{\mathsf{T}})\right)/\sqrt{2n}\right\}\right] \\
&= \mathbb{E}\left[\exp\{\beta\sqrt{2n}\langle\sum_{a=1}^{k} \boldsymbol{\sigma}^a(\boldsymbol{\sigma}^a)^{\mathsf{T}}, \boldsymbol{G}\rangle\}\right] \\
&\stackrel{(10)}{=} \exp\left\{\beta^2 \cdot n\|\sum_{a=1}^{k} \boldsymbol{\sigma}^a(\boldsymbol{\sigma}^a)^{\mathsf{T}}\|_F^2\right\} \\
&= \exp\left\{\beta^2 \cdot n\langle\sum_{a=1}^{k} \boldsymbol{\sigma}^a(\boldsymbol{\sigma}^a)^{\mathsf{T}}, \sum_{b=1}^{k} \boldsymbol{\sigma}^b(\boldsymbol{\sigma}^b)^{\mathsf{T}}\rangle\right\} \\
&= \exp\{\beta^2 n \sum_{a,b=1}^{k} \langle \boldsymbol{\sigma}^a, \boldsymbol{\sigma}^b\rangle^2\}
\end{aligned}$$

where all the steps are straightforward, while (10) is from the moment generating function of Gaussian. Plugging into the original equation above, we have

$$\begin{aligned}
\mathbb{E}[Z_n(\beta,\lambda)^k] &= \int_{(\mathbb{S}^{n-1})^{\otimes k}} \exp\{\beta n \sum_{a=1}^{k} \lambda\langle\boldsymbol{\sigma},\boldsymbol{u}\rangle^2 + \beta^2 n \sum_{a,b=1}^{k} \langle\boldsymbol{\sigma}^a,\boldsymbol{\sigma}^b\rangle^2\} \prod_{a=1}^{k} \nu_0(d\boldsymbol{\sigma}^a) \\
&\stackrel{(11)}{=} \int_{(\mathbb{S}^{n-1})^{\otimes k}} \exp\{\beta n \sum_{a=1}^{k} \lambda q_{0a}^2 + \beta^2 n \sum_{a,b=1}^{k} q_{ab}^2\} \\
&\quad \cdot \left(\int \prod_{a=1}^{k} \delta(\langle\boldsymbol{\sigma}^a,\boldsymbol{u}\rangle - q_{0a}) \prod_{a,b=1}^{k} \delta(\langle\boldsymbol{\sigma}^a,\boldsymbol{\sigma}^b\rangle - q_{ab}) \prod_{a=1}^{k} dq_{0a} \prod_{a,b=1}^{k} dq_{ab}\right) \prod_{a=1}^{k} \nu_0(d\boldsymbol{\sigma}^a) \\
&\stackrel{(12)}{\doteq} \sup_{\boldsymbol{Q}} \exp\{\beta n \sum_{a=1}^{k} \lambda q_{0a}^2 + \beta^2 n \sum_{a,b=1}^{k} q_{ab}^2\} \\
&\quad \cdot \int_{(\mathbb{S}^{n-1})^{\otimes k}} \prod_{a=1}^{k} \delta(\langle\boldsymbol{\sigma}^a,\boldsymbol{u}\rangle - q_{0a}) \prod_{a,b=1}^{k} \delta(\langle\boldsymbol{\sigma}^a,\boldsymbol{\sigma}^b\rangle - q_{ab}) \prod_{a=1}^{k} \nu_0(d\boldsymbol{\sigma}^a) \\
&\stackrel{(13)}{=} \sup_{\boldsymbol{Q}} \exp\left\{.n\left(\beta \sum_{a=1}^{k} \lambda q_{0a}^2 + \beta^2 \sum_{a,b=1}^{k} q_{ab}^2 + \frac{1}{2}\log\det(\boldsymbol{Q})\right)\right\}
\end{aligned}$$

where (11) is using the fact that:

$$1 = \int \prod_{a=1}^{k} \delta(\langle\boldsymbol{\sigma}^a,\boldsymbol{u}\rangle - q_{0a}) \prod_{a,b=1}^{k} \delta(\langle\boldsymbol{\sigma}^a,\boldsymbol{\sigma}^b\rangle - q_{ab}) \prod_{a=1}^{k} dq_{0a} \prod_{a,b=1}^{k} dq_{ab}$$

and (12) is from the Laplace method using that integration over $\boldsymbol{Q}$ is low dimensional, and (13) is direct result of the large deviation of overlap matrix from the first part of this lecture. This finishes the proof of the lemma.