Eur. Phys. J. B **29**, 457–468 (2002)
DOI: 10.1140/epjb/e2002-00326-3

THE EUROPEAN PHYSICAL JOURNAL B

On the finite size corrections to some random matching problems

G. Parisi[1,a] and M. Ratiéville[1,2,b]

[1] Dipartimento di Fisica, INFM and INFN, Università di Roma 1 *La Sapienza* P.le A. Moro, 2 – 00185 Roma, Italy
[2] Laboratoire de Physique Théorique et Modèles Statistiques Université Paris Sud, 91405 Orsay, France

Received 28 April 2002
Published online 14 October 2002 –

Abstract. We get back to the computation of the leading finite size corrections to some random link matching problems, first adressed by Mézard and Parisi [J. Phys. France **48**, 1451 (1987)]. In the so-called bipartite case, their result is in contradiction with subsequent works. We show that they made some mistakes, and correcting them, we get the expected result. In the non bipartite case, we agree with their result but push the analytical treatment further.

PACS. 75.10.Nr Spin-glass and other random models – 02.60.Pn Numerical optimization

1 Introduction

The possibility of investigating some optimization problems using techniques of the field of disordered systems in statistical physics has been recognized for a long time (see [1] for a recent review). Along with the traveling salesman problem and, more recently, K-Sat, one of the problems which got most of the attention is the matching problem, of which two variants have been studied:

(a) The simple matching problem: consider $2N$ points and a set of 'distances' between them $l_{ij} = l_{ji}$. A matching of these points is a set of N pairs so that each point belong to exactly one pair. The 'length' of such a matching is

$$L_{matching} = \sum_{pair \in matching} l_{pair}. \qquad (1)$$

One focuses on the properties of the matching of minimal length.

(b) The bipartite matching problem (or assignment problem), which is as above, except that we split the points into two sub-sets A and B of N points each, and allow only matchings where each pair is made of a point from A and a point from B.

Here we are interested in the case where the l_{ij}, $i < j$, are independent identically distributed random variables, either uniformly distributed on the interval $[0, 1]$ (so-called *flat case*) or distributed with the law $\exp(-l)$ on $[0, +\infty[$ (so-called *exponential case*).

Both the simple and bipartite cases have been investigated in the thermodynamical limit $N \to +\infty$, where self-averaging of the optimal length occurs. The replica method, in the replica symmetric scheme, yielded predictions for the mean optimal length and the distribution of the lengths of occupied links in the optimal configuration [2]. This was shown to be equivalent to a cavity approach [3]. Numerical works checked the validity of the results obtained with these techniques [4,5], and got interested in another quantity, the probability for some given point to be connected to its kth nearest neighbor in the optimal matching. This was also dealt with by an analytical cavity computation [6]. Remarkably [7] confirmed all the above results by rigorous proof.

The stability of the replica symmetric solution was checked in [8], yielding as a byproduct the $O(1/N)$ correction to the mean length of the minimal matching. For the assignment problem with flat distances, they found

$$\overline{L}^{flat}_{min} = \frac{\pi^2}{6} - \frac{1}{N}\left(\frac{\pi^2}{12} + 2\zeta(3)\right) + o\left(\frac{1}{N}\right), \qquad (2)$$

where $\overline{\cdots}$ means the average with respect to the distribution of the distances.

This seemed to agree with the numerical simulations at the time [4]. But [9] came up with a conjecture for the assignment problem at any finite N: in the exponential case the mean length of the optimal matching would be

$$\overline{L}^{exp}_{min}|_N = \sum_{k=1}^{N} \frac{1}{k^2}, \qquad (3)$$

which implies

$$\overline{L}^{exp}_{min} = \frac{\pi^2}{6} - \frac{1}{N} + o\left(\frac{1}{N}\right)\cdot \qquad (4)$$

[a] e-mail: `giorgio.parisi@roma1.infn.it`
[b] e-mail: `matthieu.ratieville@roma1.infn.it`

The problem is that in the framework of [8] it is not difficult to prove – as we shall – that

$$\overline{L}_{min}^{exp} - \overline{L}_{min}^{flat} = \frac{2\zeta(3)}{N}, \tag{5}$$

so that (4) is not compatible with (2). Beside, other results sustain (4): more recent numerical simulations [10], and an allegedly rigorous proof of (3) [11].

The main purpose of this article is to show that Mézard and Parisi [8] actually made some mistakes in the computation leading to (2). Correcting them, one does get

$$\overline{L}_{min}^{flat} = \frac{\pi^2}{6} - \frac{1}{N}\left(1 + 2\zeta(3)\right) + o\left(\frac{1}{N}\right), \tag{6}$$

which, using (5), coincides with (4).

Apart from some trivial errors due to a confusion about the number of points, N or $2N$, that has already been pointed out [4], and some typos, there are essentially two mistakes in [8]:

(i) they forgot a contribution to the $O(1/N)$ term of $\overline{L}_{min}$, but this is without any consequence as this term turns out to vanish in the zero temperature limit (Appendix A.1);

(ii) they made a mistake in the computation of another contribution (Appendix B.3), and this was responsible for the wrong result (2).

As the computations carried out in [8] are quite involved, we have chosen to make this article the most self-contained possible by restating all the necessary steps.

In Section 2, we tackle the simple matching problem, which is formally simpler but very similar to the assignment problem. Error (i) is common to both problems, and we deal with it in that section. Moreover we refine the computation of the $O(1/N)$ correction which in [8] relies on a rough numerical procedure. We give an analytical expression of the correction as the sum of a series. Unfortunately we were not able to sum this series, but it might not be impossible.

In Section 3, we turn to the assignment problem. There we correct error (ii), which is specific to this variant.

2 The non bipartite case

In this section, for the sake of simplicity, we exclusively consider the *flat case*.

To tackle the problem with the tools of statistical mechanics, one introduces an inverse temperature β (to be sent to $+\infty$ in the end) and a partition function

$$Z = \sum_{\text{all possible matchings}} \exp(-N\beta L_{matching}). \tag{7}$$

The scaling factor N ensures a good thermodynamic limit at fixed β [12].

We will not get into the details of the computation of the averaged replicated partition function, because it is quite similar to the bipartite case, for which the derivation can be found in Appendix B.1. Let us just state the result [8]:

$$\overline{Z^n} = \int \prod_\alpha \frac{\mathrm{d}Q_\alpha}{\sqrt{2\pi g_\alpha/N}} \exp\left(-\frac{N}{2}S[Q]\right) \times \exp\left(-\frac{1}{4}{\sum_{\alpha,\gamma}}' \frac{g_\alpha g_\gamma}{g_{\alpha\cup\gamma}^2} Q_{\alpha\cup\gamma}^2\right), \tag{8}$$

where $g_p = 1/(\beta p)$ and

$$S[Q] = \sum_\alpha \frac{Q_\alpha^2}{g_\alpha} - 4\ln z[Q],$$
$$z[Q] = \left(\prod_{a=1}^n \int_0^{2\pi} \frac{\mathrm{d}\lambda^a}{2\pi} \mathrm{e}^{\mathrm{i}\lambda_a}\right) \exp\left(\sum_\alpha Q_\alpha \mathrm{e}^{-\mathrm{i}\sum_{a\in\alpha}\lambda^a}\right). \tag{9}$$

In the above expressions, α stands for any non empty subset of $\{1,\dots n\}$ so that the number of Q_α variables is $2^n - 1$. For such an α, we call $p(\alpha)$ its cardinal number, and use the shorthand notation g_α for $g_{p(\alpha)}$. The notation $\sum'_{\alpha,\gamma}$ means that the summation runs over all the couples (α,γ) such that $\alpha\cap\gamma = \emptyset$.

$\overline{L}_{min}$ is nothing but the *intensive* free energy $F/N = -1/(\beta N)\ln Z$ in the limit $\beta\to+\infty$. It is evaluated by a saddle-point method. The saddle-point equations read

$$\frac{Q_\alpha}{g_\alpha} = 2\frac{\partial \ln z}{\partial Q_\alpha}\cdot \tag{10}$$

It has been solved in the limit $n\to 0$ under the assumption of replica symmetry [2]: $Q_\alpha^{sp} = Q_{p(\alpha)}^{sp}$. It turns out that the order parameters Q_p^{sp} are not well defined quantities at low temperature, and one can bypass this difficulty by considering the well defined generating function

$$G(l) = \sum_{p=1}^{+\infty} \frac{(-1)^{p-1}}{p!} Q_p^{sp} \mathrm{e}^{pl}, \tag{11}$$

for which (10) translates into

$$G(l) = -\frac{2}{\beta}\int_{-\infty}^{+\infty} \mathrm{d}y K(l+y)\mathrm{e}^{-G(y)}, \tag{12}$$

where

$$K(u) = \sum_{p=1}^{+\infty} \frac{(-1)^p}{(p!)^2}\mathrm{e}^{pu} = -1 + \mathrm{J}_0(2\mathrm{e}^{u/2}). \tag{13}$$

In (13) J_0 is the Bessel function of order zero. Note that (12) can be obtained by direct probabilistic arguments using the cavity method [3].

The free energy in the thermodynamical limit has been computed in [2]:

$$\frac{F}{N} = \lim_{\substack{n\to 0\\ \beta\to+\infty}} \frac{1}{2\beta n} S[Q_{sp}] = \frac{\pi^2}{12}\cdot \tag{14}$$

The first finite size correction ΔF is the sum of two terms: the first one, corresponding to the last factor in (8),

$$\Delta F^1 = \frac{1}{4Nn\beta} \sum_{\alpha,\gamma}{}' g_\alpha g_\gamma \left(\frac{Q^{sp}_{\alpha\cup\gamma}}{g_{\alpha\cup\gamma}}\right)^2 = -\frac{\zeta(3)}{2N} \tag{15}$$

has been computed in the appendix of [8]. The second one corresponds to the Gaussian fluctuations around the saddle-point:

$$\begin{aligned}\Delta F^2 &= \frac{1}{2\beta Nn} \ln\det\left[\frac{1}{2}\sqrt{g_\alpha}\sqrt{g_\gamma}\frac{\partial^2 S}{\partial Q_\alpha Q_\gamma}\right] \\ &= \frac{1}{2\beta Nn} \ln\det T^+_{\alpha\gamma},\end{aligned} \tag{16}$$

where we define the matrices $T^\varepsilon_{\alpha\gamma}$, $\varepsilon = \pm 1$, as follows:

$$T^\varepsilon_{\alpha\gamma} = \delta_{\alpha\gamma} + \varepsilon\frac{Q^{sp}_\alpha Q^{sp}_\gamma}{2\sqrt{g_\alpha g_\gamma}} - \varepsilon\delta_{\alpha\cap\gamma=\emptyset}\frac{Q^{sp}_{\alpha\cup\gamma}}{g_{\alpha\cup\gamma}}\sqrt{g_\alpha g_\gamma}. \tag{17}$$

At this stage the introduction of ε is a useless complication, since for the non bipartite case we only have to consider the case $\varepsilon = +1$. However the case $\varepsilon = -1$ will turn out to be useful in Section 3, where we deal with the bipartite case.

The computation of ΔF^2 involves finding the eigenvalues of $(T^\varepsilon_{\alpha\gamma})$. A vector (Q_α) is eigenvector for the eigenvalue λ if

$$\begin{aligned}\forall\alpha,\quad Q_\alpha + \varepsilon\sum_\gamma Q_\gamma \frac{Q^{sp}_\alpha Q^{sp}_\gamma}{2\sqrt{g_\alpha g_\gamma}} \\ - \varepsilon\sum_{\gamma/\alpha\cap\gamma=\emptyset} Q_\gamma\frac{Q^{sp}_{\alpha\cup\gamma}}{g_{\alpha\cup\gamma}}\sqrt{g_\alpha g_\gamma} = \lambda Q_\alpha.\end{aligned} \tag{18}$$

The diagonalization process adopted by [8] follows the de Almeida-Thouless strategy [13] of considering stable subspaces of increasing dimension. First we look for eigenvectors where no particular replica is distinguished:

$$Q_\alpha = c_{p(\alpha)}. \tag{19}$$

When we plug this into (18), we see that we have to diagonalize a $n\times n$ matrix $N^{0,\varepsilon}(n)$

$$\begin{aligned}N^{0,\varepsilon}(n)_{pq} = \delta_{pq} - \varepsilon C^q_{n-p}\frac{Q^{sp}_{p+q}}{g_{p+q}}\sqrt{g_p g_q} \\ +\varepsilon\frac{Q^{sp}_p Q^{sp}_q}{2\sqrt{g_p g_q}} C^q_n \quad p,q = 1,\dots n,\end{aligned} \tag{20}$$

where we use the notation $C^p_n = \frac{n!}{p!(n-p)!}$. This matrix turns into an infinite dimensional matrix $N^{0,\varepsilon}$ when $n \to 0$:

$$\begin{aligned}N^{0,\varepsilon}_{pq} = \delta_{pq} - \varepsilon(-1)^q\frac{(p+q-1)!}{(p-1)!q!}\frac{Q^{sp}_{p+q}}{g_{p+q}}\sqrt{g_p g_q} \\ p,q = 1,\dots +\infty\end{aligned} \tag{21}$$

(the last term in (20) does not contribute).

The eigenvalues of $N^{0,\varepsilon}$ have multiplicity 1 in the spectrum of $T^\varepsilon_{\alpha\gamma}$.

Then we look for eigenvectors of $T^\varepsilon_{\alpha\gamma}$ where one replica is distinguished, say a:

$$Q_\alpha = \begin{cases} d_{p(\alpha)} & \text{if } a \in \alpha \\ e_{p(\alpha)} & \text{if not}\end{cases}.$$

The orthogonality constraint between this family and the previous one (19) reads $pd_p + (n-p)e_p = 0$. So we can choose the only d_p as variables (of which $d_n = 0$), and we end up with the diagonalization of an $(n-1)\times(n-1)$ matrix $N^{1,\varepsilon}(n)$

$$\begin{aligned}N^{1,\varepsilon}(n)_{pq} = \delta_{pq} - \varepsilon C^q_{n-p}\frac{Q^{sp}_{p+q}}{g_{p+q}}\sqrt{g_p g_q}\frac{q}{q-n} \\ +\varepsilon C^q_{n-1}\frac{Q^{sp}_p Q^{sp}_q}{2\sqrt{g_p g_q}}\frac{q}{q-n} + \varepsilon C^{q-1}_{n-1}\frac{Q^{sp}_p Q^{sp}_q}{2\sqrt{g_p g_q}} \\ p,q = 1,\dots n-1.\end{aligned} \tag{22}$$

We get $n-1$ eigenvalues $\lambda_1,\dots\lambda_{n-1}$ (independent on a) and the corresponding eigenvectors $u^a_1,\dots u^a_{n-1}$. The important point is that the u^a_i, $a = 1,\dots n$ are not linearly independent: $\sum_a u^a_i$ is a vector of the previous family and orthogonal to it, so it has to be 0. Eventually the eigenvalues of $N^{1,\varepsilon}(n)$ have multiplicity $n-1$ in the spectrum of $T^\varepsilon_{\alpha\gamma}$. When $n\to 0$, $N^{1,\varepsilon}(n)$ becomes an infinite dimensional matrix which happens to be exactly $N^{0,\varepsilon}$.

More generally, one finds the whole spectrum of $T^\varepsilon_{\alpha\gamma}$ by looking for eigenvectors which have k given distinguished replicas:

$$Q_\alpha = \begin{cases} 0 & \text{if } p_\alpha < k \\ d^i_{p(\alpha)} & \text{if } \alpha \text{ contains } k+1-i \text{ of the distinguished replicas,}\end{cases}$$

where i goes from one to $k+1$. The orthogonalization with respect to a family where only $k-1$ of these replicas are distinguished reads as a system of equations

$$\forall j = 0,1,\dots k-1, \qquad \sum_{r=0}^{k-j} C^r_{k-j} C^{p-(r+j)}_{n-k} d^{k+1-(r+j)}_p = 0, \tag{23}$$

whose solution in the $n\to 0$ limit is

$$\begin{aligned}\frac{d^1_p}{p(p+1)\dots(p+k-1)} &= \frac{d^2_p}{(p-k+1)(p+1)\dots(p+k-1)} \\ &= \frac{d^3_p}{(p-k+1)(p-k+2)(p+2)\dots(p+k-1)} \\ &= \dots = \frac{d^{k+1}_p}{(p-k+1)\dots p}\end{aligned} \tag{24}$$

(note that this is slightly different from equation (20) in [8] where there is a typo)

It follows that we can keep the only d^1_p as independent variables, and have to diagonalize a matrix $N^{k,\varepsilon}(n)$, which in the limit $n \to 0$ is the infinite dimensional matrix

$$N^{k,\varepsilon}_{pq} = \delta_{p,q} - \varepsilon(-1)^q \frac{(p+q-1)!(q-1)!}{(p-1)!(q-k)!(q+k-1)!}\frac{Q^{sp}_{p+q}}{g_{p+q}}\sqrt{g_p g_q}$$
$$p, q = 1, \ldots + \infty. \quad (25)$$

The eigenvalues of $N^{k,\varepsilon}(n)$ have multiplicity $C^k_n - C^{k-1}_n$ in the spectrum of $T^{\varepsilon}_{\alpha\gamma}$.

The article [8] prefers using matrices derived from the $N^{k,\varepsilon}$ by some transformations which do not affect the spectrum: after shifting the indices p and q by k, then transposing and multiplying each entry by $(-1)^{p+q}\sqrt{\frac{g_{q+k}}{g_{p+k}}}\frac{(q+1)\ldots(q+k-1)}{(p+1)\ldots(p+k-1)}$, one gets the family of matrices

$$M^{k,\varepsilon}_{p,q} = \delta_{p,q} - \varepsilon(-1)^{q+k}\frac{(p+q+2k-1)!}{(p+2k-1)!q!}g_{q+k}\frac{Q^{sp}_{p+q+2k}}{g_{p+q+2k}}, \quad (26)$$

where $p, q = 0, 1, \ldots + \infty$.

We can now proceed to the computation of ΔF^2 (16):

$$\Delta F^2 = \Delta F^{2,+}_1 + \Delta F^{2,+}_2, \quad (27)$$

where

$$\Delta F^{2,\varepsilon}_1 = \lim_{n\to 0}\frac{1}{2\beta n N}\left[\ln\det N^{0,\varepsilon}(n) + (n-1)\ln\det N^{1,\varepsilon}(n)\right],$$
$$\Delta F^{2,\varepsilon}_2 = \lim_{n\to 0}\frac{1}{2\beta n N}\sum_{k\geq 2}\left(C^k_n - C^{k-1}_n\right)\ln\det N^{k,\varepsilon}(n). \quad (28)$$

For $k \geq 2$ one has $\left(C^k_n - C^{k-1}_n\right) \sim n(-1)^{k-1}\frac{2k-1}{k(k-1)}$ so that

$$\Delta F^{2,\varepsilon}_2 = \frac{1}{2\beta N}\sum_{k=2}^{+\infty}(-1)^{k-1}\frac{2k-1}{k(k-1)}\ln\det M^{k,\varepsilon}. \quad (29)$$

There is a subtlety in the computation of $\Delta F^{2,+}_1$: as the limits of $N^{0,+}(n)$ and $N^{1,+}(n)$ when $n \to 0$ are the same, one may be tempted to say that in this limit we have a unique family of eigenvalues of multiplicity n, and so $\Delta F^{2,+}_1 = 1/(2\beta N)\ln\det M^{1,+}$. It is what [8] did, but it is wrong. Actually there is a factor $1/n$ to take into account, so that one also gets the contribution of the derivatives

$$\Delta F^{2,+}_1 = \frac{1}{2\beta N}\left[\left\{\frac{\mathrm{d}\ln\det N^{0,+}(n)}{\mathrm{d}n} - \frac{\mathrm{d}\ln\det N^{1,+}(n)}{\mathrm{d}n}\right\}_{n=0} + \ln\det M^{1,+}\right]. \quad (30)$$

It turns out that the extra term is zero when $\beta \to +\infty$, but it is not trivial (see Appendix A.1). We also show *en passant* in this appendix that $\det M^{1,+}$ has a non zero finite limit when $\beta \to +\infty$ (1 actually) so that we eventually agree with [8] on the fact that

$$\Delta F^{2,+}_1 = 0. \quad (31)$$

As far as the computation of ΔF^2_2 (29) is concerned, the strategy of [8] consists into translating the infinite dimensional matrices $M^{k,\varepsilon}$ into more tractable integral operators. If (c_p) is an eigenvector of $M^{k,\varepsilon}$, then

$$f(x) = \sum_{q=0}^{+\infty}\frac{(-1)^q}{q!}\sqrt{g_{q+k}}c_q \mathrm{e}^{(k+q)x - G(x)/2} \quad (32)$$

is an eigenfunction, with the same eigenvalue, of the operator

$$M^{k,\varepsilon}(x,y) = \delta(x-y) - \varepsilon(-1)^k A^k(x,y), \quad (33)$$

where

$$A^k(x,y) = 2\exp\left(-\frac{G(x)+G(y)}{2} + k(x+y)\right)\sum_{p=0}^{+\infty}\frac{(-1)^p \mathrm{e}^{p(x+y)}}{p!(2k+p-1)!}g_{p+k} \quad (34)$$

and reciprocally.

The article [8], on the basis of numerical discretization and diagonalization of these operators, argues that the values of $\det M^{k,\varepsilon}(T)$ plotted *versus* $T\ln k$ fall onto two universal curves, depending on the parity of k:

$$\det M^{k,\varepsilon}(T) = \begin{cases} f_\varepsilon(T\ln k) & \text{if } k \text{ is even} \\ f_{-\varepsilon}(T\ln k) & \text{if } k \text{ is odd} \end{cases}. \quad (35)$$

A more accurate numerical analysis showed us that this happens only when $k \to +\infty$, $T\ln k$ being kept fixed: f_+ and f_- are limit functions (see Fig. 1). Happily this does not change the conclusion that, in the limit $\beta \to +\infty$,

$$\Delta F^{2,\varepsilon}_2 = -\frac{1}{2N}\int_0^{+\infty}\mathrm{d}t[\ln f_\varepsilon(t) - \ln f_{-\varepsilon}(t)]. \quad (36)$$

In [8], this integral is performed by fitting the numerical curves of f_+ and f_- by smooth functions, which yields an estimate flawed by a rather rough uncertainty:

$$\Delta F^{2,+}_2 = \frac{1}{N}(0.47 \pm 0.05). \quad (37)$$

As a consequence, ΔF itself is known with a bad precision:

$$\Delta F = \frac{1}{N}(-0.13 \pm -0.05). \quad (38)$$

It is possible to improve on this. One can explicitly compute the limit of the operator $A^k(x,y)$ when $k \to \infty$ under the restriction that $t = T\ln k$ remains fixed. In this case one also has $\beta \to +\infty$. Let us recall that in this limit [2]

$$G(l) = \hat{G}(\beta l) \text{ where } \hat{G}(x) = \ln(1+\mathrm{e}^{2x}). \quad (39)$$

So we can write

$$A^k(x,y) \sim \frac{2}{\beta\sqrt{(1+\exp(2x/\beta))(1+\exp(2y/\beta))}} \times f(k, \exp(x+y)), \quad (40)$$

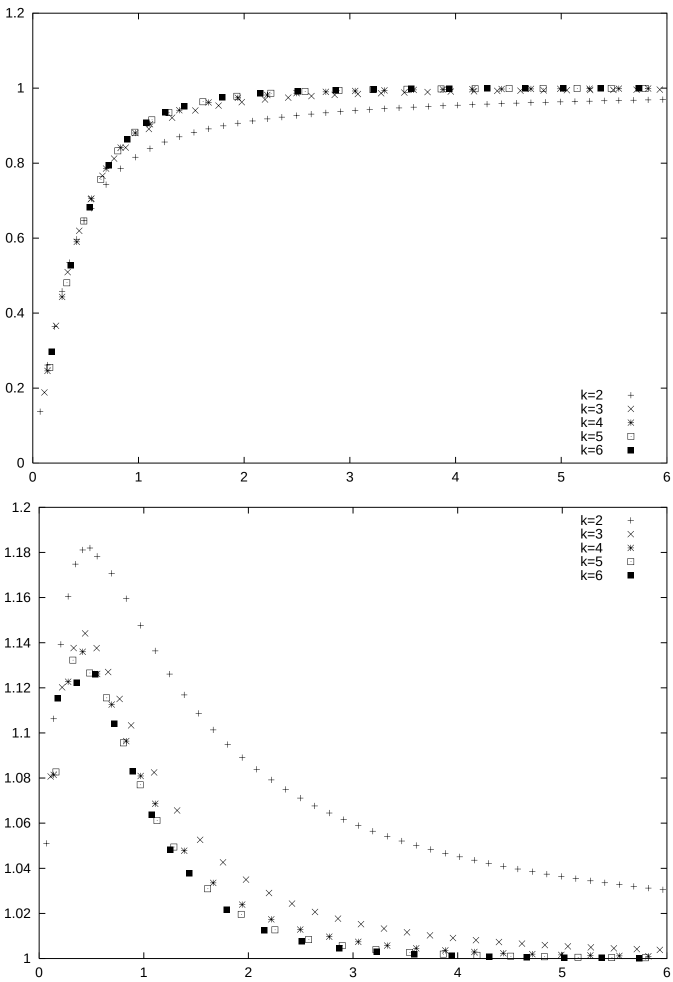

Fig. 1. The values of $\det M^{k,\varepsilon}(T)$ defined in (26) plotted *versus* $T \ln k$. Above, $\det M^{k,-}$ k odd and $\det M^{k,+}$ k even. Below, $\det M^{k,-}$ k even and $\det M^{k,+}$ k odd.

where

$$f(k,z) = z^k \sum_{p=0}^{+\infty} \frac{(-1)^p z^p}{p!(2k+p-1)!} \frac{1}{p+k} \cdot \tag{41}$$

The eigenvalues of $A^k(x,y)$ are the same as the ones of the operator

$$\frac{\ln k}{t} A^k \left(\frac{\ln k}{t} u, \frac{\ln k}{t} v \right) = \frac{2}{\sqrt{(1+\exp(2u))(1+\exp(2v))}} \times f\left(k, \exp\left(\frac{\ln k}{t}(u+v)\right)\right). \tag{42}$$

Let us define

$$g(k,w) = f(k, \exp(w \ln k)). \tag{43}$$

In Appendix A.2 we show that when $k \to +\infty$, $g(k,w) \to \Theta(w-2)$ where Θ is the usual Heaviside step function. So the operator we have to diagonalize is

$$\frac{2}{\sqrt{(1+\exp(2u))(1+\exp(2v))}} \Theta\left(\frac{u+v}{t} - 2\right). \tag{44}$$

It is the same as diagonalizing

$$H_t(u,v) = \frac{2}{\sqrt{(1+\exp(2(u+t)))(1+\exp(2(v+t)))}} \Theta(u+v). \tag{45}$$

We have $f_+(t) = \ln\det(I - H_t)$ and $f_-(t) = \ln\det(I + H_t)$. The correction (36) reads

$$\Delta F_2^{2,+} = \frac{1}{2N} \int_0^{+\infty} \mathrm{d}t \, [\ln(\det(I+H_t)) - \ln(\det(I-H_t))] \tag{46}$$

$$= \frac{1}{N} \sum_{p=0}^{+\infty} \frac{I_{2p+1}}{2p+1} \tag{47}$$

where

$$\begin{aligned} I_p &= \int_0^{+\infty} \mathrm{d}t \, \mathrm{Tr} H_t^p = \int_0^{+\infty} \mathrm{d}t \\ &\quad \times \int \mathrm{d}u_1 \ldots \mathrm{d}u_p H_t(u_1,u_2) \ldots H_t(u_{p-1},u_p) H_t(u_p,u_1) \\ &= 2^p \int_0^{+\infty} \mathrm{d}t \int \mathrm{d}u_1 \ldots \mathrm{d}u_p \\ &\quad \times \frac{\Theta(u_1+u_2) \ldots \Theta(u_{p-1}+u_p) \Theta(u_p+u_1)}{(1+\exp(2(t+u_1))) \ldots (1+\exp(2(t+u_p)))} \end{aligned} \tag{48}$$

(to derive (46) we used the identity $\ln\det = \mathrm{Tr}\, \ln$ and expanded $\ln(\det(I \pm H_t))$ in power series of H_t).

Note that the operators H_t have positive and negative eigenvalues. The changes of variables $x_i = \exp(-2u_i)$ and $C = \exp(2t)$ yield

$$I_p = \int_0^{+\infty} \mathrm{d}t \, Tr H_t^p = \int_1^{+\infty} \frac{\mathrm{d}C}{2C} \int \frac{\mathrm{d}x_1}{x_1+C} \cdots \frac{\mathrm{d}x_p}{x_p+C}, \tag{49}$$

where the integration with respect to $x_1, \ldots x_p$ is to be performed over the domain defined by $\forall i, x_i \geq 0$ and $x_i x_{i+1} \leq 1$, $x_p x_1 \leq 1$.

Unfortunately we were not able to compute analytically I_p for a generic p. We succeeded in computing exactly the four first terms, and we got an estimate of the fifth one by numerical integration:

$$\begin{aligned} I_1 &= \frac{\zeta(2)}{4} \sim 0.411234, \\ I_2 &= \frac{\zeta(3)}{2} \sim 0.601028, \\ I_3 &= \frac{3\zeta(4)}{16} \sim 0.202936, \\ I_4 &= 4\zeta(5) - \frac{\pi^2 \zeta(3)}{3} \sim 0.193102, \\ I_5 &\sim 0.137098. \end{aligned} \tag{50}$$

A truncated summation of (46) up to the third term gives the following lower bound

$$\Delta F_2^{2,+} > 0.506298/N, \tag{51}$$

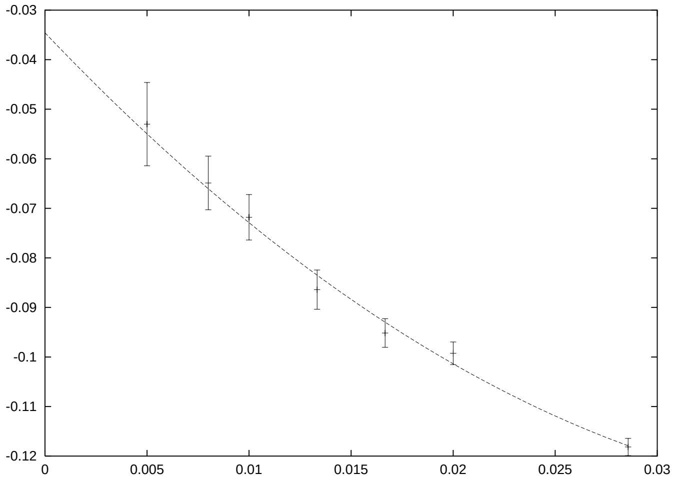

Fig. 2. $(\overline{L}_{min} - \pi^2/12)N$ *versus* $1/N$. The dashed line is the fit (56).

Adding ΔF_1 (15), we get

$$\Delta F > -0.0947301/N. \tag{52}$$

One can also try a Padé summation involving I_1, I_2 and I_3 to compute $\Delta F_2^{2,+}$. One gets

$$N\Delta F \sim I_1 + \frac{I_3/3}{1 - \frac{I_5/5}{I_3/3}} - \frac{\zeta(3)}{2} = -0.076. \tag{53}$$

To obtain a more accurate estimate of $\Delta F_2^{2,+}$, one can get back to equation (45) and (46), and use the same kind of discretization scheme that Mézard and Parisi used to get (37). We restricted the function $H_t(u,v)$ to the square $[-10,+10]^2$, outside of which it is smaller than 10^{-4}. Starting from $t = 0$, then incrementing t by 0.01, we sampled H_t on P equidistant points in each direction u and v, diagonalized the resulting $P \times P$ matrix, and computed $\det(I \pm H_t)$. These values were carefully extrapolated to $P \to +\infty$. With increasing t both values go to 1: we stopped at $t = 5$, where their differences to 1 is smaller than 10^{-4}. We eventually discretized equation (46) to get

$$N\Delta F_2^{2,+} = 0.5667 \pm 5.10^{-4}, \tag{54}$$

and so

$$N\Delta F = -0.0343 \pm 5.10^{-4}, \tag{55}$$

which is compatible with (52).

To check the validity of (55), we carried out numerical simulations similar to the ones in [4], but averaging over more samples and implementing a variance reduction trick [18,19]. We used the values $N = 35$, 50, 60, 75, 100, 125 and 200, with a decreasing number of samples, from 1 200 000 down to 300 000. The results for $\overline{L}_{min}$ are plotted in Figure 2. A quadratic fit

$$\left(\overline{L}_{min} - \frac{\pi^2}{12}\right) N = a + \frac{b}{N} + \frac{c}{N^2} \tag{56}$$

gives $a = -0.0346 \pm 0.0066$, which is in very good agreement with (55).

3 The bipartite case

Now we turn to the assignment problem. To make the comparison with the non bipartite case easier, we prefer using a slightly different convention for the partition function, which amounts to a rescaling of β: we set

$$Z = \sum_{\text{all possible matchings}} \exp\left(-\frac{N}{2}\beta L_{matching}\right) \tag{57}$$

instead of (7). However *the reader must bear in mind that in this case $\overline{L}_{min}$ is TWICE the free energy density $F/N = -1/(\beta N)\ln Z$.*

Moreover, as we want to compare our results with (4), we must take into account both possible distributions of the distances. In Appendix B.1 we sketch the derivation of the averaged replicated partition function both in the case of the flat distribution ($\mu = 0$) and in the case of the exponential distribution ($\mu = 1$):

$$\begin{aligned}\overline{Z^n} = \int \prod_\alpha \mathrm{d}X_\alpha dY_\alpha \frac{N}{2\pi g_\alpha} \exp\left(-\frac{N}{2} S[X_\alpha, Y_\alpha]\right) \\ \times \exp\left(-\frac{1}{2}\sum_{\alpha,\beta}{}' \frac{g_\alpha g_\beta}{g^2_{\alpha\cup\beta}}\left(X^2_{\alpha\cup\beta} + Y^2_{\alpha\cup\beta}\right)\right) \\ \times \exp\left(-\mu\sum_\alpha \left(X_\alpha^2 + Y_\alpha^2\right)\right),\end{aligned} \tag{58}$$

where

$$\begin{aligned}S[X_\alpha, Y_\alpha] = \sum_\alpha \frac{X_\alpha^2 + Y_\alpha^2}{g_\alpha} - 2\ln z[X_\alpha - \mathrm{i}Y_\alpha] \\ -2\ln z[X_\alpha + \mathrm{i}Y_\alpha].\end{aligned} \tag{59}$$

The thermodynamical limit of F/N does not depend on μ, but its correction in $1/N$ does. One can look for a saddle-point of the particular form:

$$X_\alpha^{sp} = X_{p(\alpha)}^{sp} \text{ and } Y_\alpha^{sp} = 0. \tag{60}$$

X^{sp} satisfy the equation

$$X_\alpha = 2g_\alpha \frac{\partial \ln z}{\partial X_\alpha}, \tag{61}$$

which is exactly the same as the one for the non bipartite case (10). Hence

$$X_p^{sp} = Q_p^{sp}. \tag{62}$$

The free energy in the thermodynamical limit is the same as in the non bipartite case (14).

Like in the non bipartite case, the $O(1/N)$ correction to the free energy contains the terms ΔF^1 coming from the last line in (58), and ΔF^2 coming from the Gaussian fluctuations.

One has $\Delta F^1 = \Delta F_1^1 + \mu \Delta F_2^1$, with

$$\Delta F_1^1 = \frac{1}{2nN\beta}\sum_{\alpha,\gamma}{}' \frac{g_\alpha g_\gamma}{g^2_{\alpha\cup\gamma}}\left(Q^{sp}_{\alpha\cup\gamma}\right)^2 \tag{63}$$

and

$$\Delta F_2^1 = \frac{1}{nN\beta}\sum_\alpha (Q_\alpha^{sp})^2 . \tag{64}$$

We see that ΔF_1^1 is twice the one computed in the case of the simple matching problem (15). ΔF_2^1, computed in Appendix B.2, happens to be the opposite of ΔF_1^1. So

$$\Delta F^1 = (\mu - 1)\frac{\zeta(3)}{N}, \tag{65}$$

which demonstrates our assertion (5), confirmed by numerical simulations [10].

As far as ΔF^2 is concerned, it is easy to generalize the computation of the non bipartite case. We have

$$\Delta F^2 = \frac{1}{2\beta N n}\ln\left[\det T_{\alpha\gamma}^{+}\det T_{\alpha\gamma}^{-}\right], \tag{66}$$

where $T_{\alpha\gamma}^{\pm}$ are the matrices of equation (17).

Thus we can write $\Delta F^2 = \Delta F_1^{2,+} + \Delta F_2^{2,+} + \Delta F_1^{2,-} + \Delta F_2^{2,-}$, where the different contributions are defined in (28). We have $\Delta F_2^{2,+} + \Delta F_2^{2,-} = 0$ (see Eq. (36)) and we know that $\Delta F_1^{2,+} = 0$ (31).

There is a subtlety however in the computation of $\Delta F_1^{2,-}$, as pointed out in [8]: we run into a problem because $T_{\alpha\beta}^{-}$ has some zero modes. This is actually no surprise: it arises from the fact that the action (59) is left invariant under the transformation

$$X_\alpha + \mathrm{i}Y_\alpha \to (X_\alpha + \mathrm{i}Y_\alpha)\exp\left(\mathrm{i}\sum_{a\in\alpha}\theta_a\right), \tag{67}$$

where $\theta_1, \ldots \theta_n$ are real angles. The zero modes are the n Goldstone modes of this invariance. As a consequence the saddle-point (60) is not unique: there is a n dimensional hypersurface of degenerated saddle-points parameterized as follows

$$X_\alpha^{sp} + \mathrm{i}Y_\alpha^{sp} = X_{p(\alpha)}^{sp}\exp\left(\mathrm{i}\sum_{a\in\alpha}\theta_a\right) \qquad 0 \le \theta_i \le 2\pi. \tag{68}$$

The kernel of $T_{\alpha\beta}^{-}$ is spanned by the n vectors ξ_i of components

$$\xi_i^\alpha = \frac{\partial Y_\alpha^{sp}}{\partial\theta_i} = \begin{cases} X_{p(\alpha)}^{sp} & \text{if } i\in\alpha \\ 0 & \text{if not}\end{cases}. \tag{69}$$

which have the replica i distinguished. So $N^{k,-}(n)$ for $k \ge 2$ has no zero mode. Only $N^{0,-}(n)$ and $N^{1,-}(n)$ have a zero eigenvalue each. Thus $\Delta F_1^{2,-}$ is to be computed leaving aside the zero modes, and one has to take into account a new contribution ΔF^3 to the free energy corresponding to the volume of the orbit (68).

As far a the computation of $\Delta F_1^{2,-}$ is concerned, we refer to Appendix A.1, where we showed that $\Delta F_1^{2,+} = 0$. It is easy to see that the proof is exactly the same for $\Delta F_1^{2,-}$: there are only some sign reversals (in particular one has $1 + \lambda_k$ instead of $1 - \lambda_k$ in the denominator of (82)), and the exclusion of the zero modes (corresponding to $k = 1$ in (87)) has no consequence because the key property (83) holds for each eigenvalue of $I(x, y)$. So $\Delta F_1^{2,-} = 0$, and

$$\Delta F^2 = 0. \tag{70}$$

Let us now turn to the computation of the volume of the hypersurface defined by (68). It is where Mézard and Parisi made a mistake: they computed this quantity without taking into account the fact that they had carried out the diagonalization in another system of coordinates. To make things clearer let us rewrite (58) as

$$\begin{aligned}\overline{Z^n} = \int \prod \mathrm{d}U_\alpha dV_\alpha \frac{N}{2\pi}\exp\left(-\frac{N}{2}S\left[\sqrt{g_\alpha}U_\alpha, \sqrt{g_\alpha}V_\alpha\right]\right) \\ \times\exp\left(-\frac{1}{2}\sum_{\alpha,\gamma}{}' \frac{g_\alpha g_\gamma}{g_{\alpha\cup\gamma}}\left(U_{\alpha\cup\gamma}^2 + V_{\alpha\cup\gamma}^2\right)\right) \\ \times\exp\left(-\mu\sum_\alpha g_\alpha\left(U_\alpha^2 + V_\alpha^2\right)\right).\end{aligned} \tag{71}$$

It is in the variables (U_α, V_α) that Mézard and Parisi have chosen to diagonalize: indeed the matrix $T_{\alpha\gamma}^{\pm}$ (17) is half the Hessian matrix of $S[\sqrt{g_\alpha}U_\alpha, \sqrt{g_\alpha}V_\alpha]$. So the volume of the hypersurface is to be computed in these same variables, not in (X_α, Y_α) as they did. The saddle-points coordinates are related by

$$\begin{cases} U_\alpha^{sp} = X_\alpha^{sp}/\sqrt{g_p} \\ V_\alpha^{sp} = Y_\alpha^{sp}/\sqrt{g_p}\end{cases}. \tag{72}$$

The computation of the correct volume is done in Appendix B.3. We find

$$\Delta F^3 = -\frac{1}{2N}\cdot \tag{73}$$

Collecting the pieces (65, 70, 73), we get the expected result (4).

4 Conclusion

By fixing the mistakes made by [8] in the computation of the $O(1/N)$ correction to the mean minimum length in the random assignment problem, we removed any inconsistency among the corpus of results about this problem. This gives further evidence that the replica approach, in its simplest symmetric ansatz, exactly solves the problem, and remains a valuable tool to gain insight on such quantities as this finite size correction which, despite recent dramatic progresses in the rigorous approach [7], still resist a mathematical treatment.

One may hope to give further support to the conjecture (3) by computing higher order terms in the expansion of $\overline{L}_{min}^{exp}$ in powers of $1/N$ when $N \to +\infty$. As far as the second order term is concerned, it is easy to see that (3) predicts

$$\overline{L}_{min}^{exp} = \frac{\pi^2}{6} - \frac{1}{N} + \frac{1}{2N^2} + o\left(\frac{1}{N^2}\right). \qquad (74)$$

In theory this extra term can be recovered by the saddle-point method: when computing the averaged replicated partition function (Appendix B.1), one must add one term to the expansion (116). Then one must work out the resulting $\overline{Z^n}$ with standard techniques [20]. We just had a glance at it, but it seems to be quite an involved program.

We acknowledge very useful discussions with A. Cavagna, J. Houdayer and M. Mézard.

Appendix A: The non bipartite case

A.1 Contribution of the derivatives in $\Delta \mathsf{F}_1^2$ (30)

Our purpose here is to show that

$$\Gamma = \lim_{\beta\to+\infty} \frac{1}{2\beta N}\left[\frac{\mathrm{d}\ln\det N^{0,+}(n)}{\mathrm{d}n} - \frac{\mathrm{d}\ln\det N^{1,+}(n)}{\mathrm{d}n}\right]_{n=0} \qquad (75)$$

equals 0.

We start from the equations (20) and (22). Writing $\frac{q}{q-n} = 1 + \frac{n}{q} + o(n)$ it is easy to see that $N^{1,+}(n) - N^{0,+}(n) = n\Delta^1 + n\Delta^2 + o(n)$, where Δ_1 and Δ_2 are the following infinite dimensional matrices

$$\Delta^1_{pq} = \frac{(-1)^{q+1}}{q}\frac{(p+q-1)!}{(p-1)!q!}\frac{Q_{p+q}}{g_{p+q}}\sqrt{g_p g_q} \qquad (76)$$

$$\Delta^2_{pq} = \frac{(-1)^q}{q}\frac{Q_p^{sp}Q_q^{sp}}{2\sqrt{g_p g_q}} \qquad p,q = 1,\ldots+\infty. \qquad (77)$$

It is more convenient to use an integral operator formalism. We set

$$I(x,y) = \frac{2}{\beta}\exp\left(-\frac{G(x)+G(y)}{2}\right)K(x+y), \qquad (78)$$

where K is the function defined in (13).

If (c_p) is an eigenvector of $N^{0,+}_{pq}$ (21) for the eigenvalue $1-\lambda$ then f is an eigenfunction of $I(x,y)$ for the eigenvalue λ, where

$$f(x) = \sum_{q=1}^{+\infty}\frac{(-1)^q}{q!}\sqrt{g_q}c_q \mathrm{e}^{qx - G(x)/2}. \qquad (79)$$

One can easily check that the operators corresponding to the matrices Δ^1 and Δ^2 are respectively

$$\Delta^1(x,y) = -\frac{2}{\beta}\mathrm{e}^{-G(x)/2}\mathrm{e}^{G(y)/2}\int_y^{+\infty}K(t+x)\mathrm{e}^{-G(t)}\mathrm{d}t, \qquad (80)$$

$$\Delta^2(x,y) = -G(x)\mathrm{e}^{-G(y)/2}\mathrm{e}^{-G(x)/2}. \qquad (81)$$

Thus (75) can be evaluated by the standard result of first order perturbation theory

$$\Gamma = -\frac{1}{2N}\sum_k \frac{\langle k|\Delta^1|k\rangle + \langle k|\Delta^2|k\rangle}{\beta(1-\lambda_k)}, \qquad (82)$$

where the $|k\rangle$ are the normalized eigenvectors of $I(x,y)$ (78), and the λ_k are the corresponding eigenvalues.

Below we show that

$$\frac{\langle k|\Delta^1|k\rangle + \langle k|\Delta^2|k\rangle}{\beta} = 0 \qquad (83)$$

for each k when $\beta\to+\infty$, and that the λ_k have finite limits, different from 1, so that $\Gamma = 0$.

Let us consider an eigenfunction f of $I(x,y)$:

$$\int \mathrm{d}y\, I(x,y)f(y) = \lambda f(x). \qquad (84)$$

We make the substitution

$$f(x)\mathrm{e}^{G(x)/2} = P\left(\frac{1}{1+\exp(2x/\beta)}\right). \qquad (85)$$

In the $\beta\to+\infty$ limit we can use (39) and after some changes of variables we see that (84) can be restated as

$$\forall v\in[0,1],\quad \int_0^1 \mathrm{d}u\frac{P(u)}{1-u}K\left[\frac{\beta}{2}\ln\left(\frac{(1-u)(1-v)}{uv}\right)\right] = \lambda P(v),$$

$$i.e.\ -\int_{1-v}^1 \mathrm{d}u\frac{P(u)}{1-u} = \lambda P(v), \qquad (86)$$

because $K(\beta z) = -1$ if $z\geq 0$, 0 otherwise. Note that $P(0) = 0$.

The eigenproblem (86) is β independent. We found that its eigenvalues are $\lambda_k = (-1)^k/k$, $k = 1,2,\ldots+\infty$. The corresponding eigenfunctions are polynomial of degree k:

$$P_k(u) = \sum_{p=0}^k a_{p,k}u^p \quad \text{with}$$

$$a_{p,k} = (-1)^p\frac{k^2(k^2-1)\ldots(k^2-(p-1)^2)}{(p!)^2}. \qquad (87)$$

For the computation of $\langle k|\Delta^1|k\rangle$, it is simpler not to use this explicit form. By derivation of (86), we get

$$P(1-v) = \lambda v P'(v). \qquad (88)$$

Combining (78, 80) and (84), we have

$$\int \Delta^1(x,y)f(x)\mathrm{d}x = -\lambda\mathrm{e}^{G(y)/2}\int_y^{+\infty}f(t)\mathrm{e}^{-G(t)/2}, \qquad (89)$$

so that

$$\int\!\!\int \Delta^1(x,y)f(x)f(y)\mathrm{d}x\mathrm{d}y =$$

$$\lambda\frac{\beta^2}{4}\int_0^1\mathrm{d}u\frac{P(u)}{u(1-u)}\int_u^1\mathrm{d}v\frac{P(v)}{1-v}$$

$$= -\lambda^2\frac{\beta^2}{4}\int_0^1\mathrm{d}u\frac{P(u)}{u(1-u)}P(1-u) \quad \text{by (88)}. \qquad (90)$$

This can be further simplified:

$$\int_0^1 \mathrm{d}u \frac{P(u)P(1-u)}{u(1-u)} = \int_0^1 \mathrm{d}u P(u)P(1-u)\left(\frac{1}{u}+\frac{1}{1-u}\right)$$
$$= 2\int_0^1 \mathrm{d}u \frac{P(u)P(1-u)}{u}$$
$$= 2\lambda \int_0^1 \mathrm{d}u\, P(u)P'(u)$$
$$= \lambda P(1)^2, \qquad (91)$$

so that

$$\int\int \Delta^1(x,y)f(x)f(y)\mathrm{d}x\mathrm{d}y = -\lambda^3 \frac{\beta^2}{4}P(1)^2. \qquad (92)$$

The function f is *a priori* not normalized so that the above quantity is to be divided by

$$\int f(x)^2\mathrm{d}x = \frac{\beta}{2}\int_0^1 \mathrm{d}u \frac{P(u)^2}{1-u}\cdot \qquad (93)$$

We eventually get

$$\langle k|\Delta^1|k\rangle = -\lambda^3\frac{\beta}{2}P(1)^2\left[\int_0^1 \mathrm{d}u\frac{P(u)^2}{1-u}\right]^{-1}\cdot \qquad (94)$$

Now we compute $\langle k|\Delta^2|k\rangle$: thanks to (12),

$$\Delta^2(x,y) = \frac{2}{\beta}\int \mathrm{d}t K(t+x)\mathrm{e}^{-G(t)}\mathrm{e}^{-G(x)/2}\mathrm{e}^{-G(y)/2}$$
$$= \int \mathrm{d}t I(x,t)\mathrm{e}^{-G(y)/2}\mathrm{e}^{-G(t)/2}, \qquad (95)$$

hence

$$\int \Delta^2(x,y)f(x)\mathrm{d}x = \lambda\int \mathrm{d}t f(t)\mathrm{e}^{-G(y)/2}\mathrm{e}^{-G(t)/2}, \qquad (96)$$

$$\int\int \Delta^2(x,y)f(x)f(y)\mathrm{d}x\mathrm{d}y = \lambda\left[\int \mathrm{d}t f(t)\mathrm{e}^{-G(t)/2}\right]^2$$
$$= \lambda\left[\frac{\beta}{2}\int_0^1 \mathrm{d}u\frac{P(u)}{1-u}\right]^2$$
$$= \lambda^3\frac{\beta^2}{4}P(1)^2 \qquad (97)$$

and $\langle k|\Delta^2|k\rangle = +\lambda^3\frac{\beta}{2}P(1)^2\left[\int_0^1 \mathrm{d}u\frac{P(u)^2}{1-u}\right]^{-1}$, which ends up the proof.

A.2 The limit of g(k, w) when k → +∞

We start from an integral representation of $g(k,w)$ defined in (43):

$$g(k,w) = \mathrm{e}^{kw\ln k}\sum_{p=0}^{+\infty}\frac{(-1)^p\mathrm{e}^{pw\ln k}}{p!(p+2k-1)!}\frac{1}{p+k}$$
$$= \frac{\mathrm{i}}{2\pi}\mathrm{e}^{kw\ln k}\int_C \mathrm{d}z\int_0^{+\infty}\mathrm{d}x\mathrm{e}^{S(k,z,x)}, \qquad (98)$$

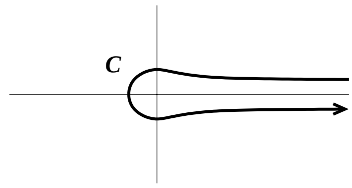

Fig. 3. The contour C in the complex plane.

where

$$S(k,z,x) = -z - k(x+2\ln(-z)) + \frac{k^w}{z}\mathrm{e}^{-x}, \qquad (99)$$

because $(p+k)^{-1} = \int_0^{+\infty}\mathrm{d}x\,\mathrm{e}^{-(p+k)x}$ and $(p+2k-1)!^{-1} = \mathrm{i}/(2\pi)\int_C \mathrm{d}z\,\mathrm{e}^{-z}(-z)^{-(p+2k)}$. C is a contour in the complex plane such as illustrated in Figure 3.

The stationarity equations read

$$\frac{\partial S}{\partial z} = -1 - 2\frac{k}{z} - \frac{k^w}{z^2}\mathrm{e}^{-x} = 0,$$
$$\frac{\partial S}{\partial x} = -k - \frac{k^w}{z}\mathrm{e}^{-x} = 0, \qquad (100)$$

so that there is a movable saddle-point at

$$\begin{cases} x_{sp} = (w-2)\ln k \\ z_{sp} = -k \end{cases}. \qquad (101)$$

A.2.1 The case w > 2

In this case the saddle-point (101) is inside the range of integration. We compute the Hessian at this point:

$$\frac{\partial^2 S}{\partial t^2} = 2\frac{k}{z^2} + 2\frac{k^w}{z^3}\mathrm{e}^{-x} = 0,$$
$$\frac{\partial^2 S}{\partial x^2} = \frac{k^w}{z}\mathrm{e}^{-x} = -k,$$
$$\frac{\partial^2 S}{\partial x\partial z} = \frac{k^w}{z^2}\mathrm{e}^{-x} = 1. \qquad (102)$$

So, when $k\to+\infty$,

$$\int_0^{+\infty}\mathrm{d}x\mathrm{e}^{S[k,z,x]} \sim \int_0^{+\infty}\mathrm{d}x$$
$$\times \exp\left[S_{sp} - \frac{k}{2}[x-(w-2)\ln k]^2 + [x-(w-2)\ln k](z+k)\right]$$
$$\sim \sqrt{\frac{2\pi}{k}}\mathrm{e}^{S_{sp}}\exp\left[\frac{1}{2k}(z+k)^2\right]. \qquad (103)$$

Then we perform the integration with respect to z with $z+k = -\mathrm{i}\epsilon$:

$$\int \mathrm{d}z\mathrm{e}^{\frac{1}{2k}(z+k)^2} = -\mathrm{i}\int_{-\infty}^{+\infty}\mathrm{d}\epsilon\mathrm{e}^{-\frac{\epsilon^2}{2k}}$$
$$= -\mathrm{i}\sqrt{2\pi k}. \qquad (104)$$

Given that $S_{sp} = -kw \ln k$, we eventually get

$$\int\int \mathrm{d}x\mathrm{d}z \mathrm{e}^{S(k,z,x)} \sim -2\pi \mathrm{i} \mathrm{e}^{-kw \ln k}. \tag{105}$$

It follows that

$$g(k, w) \to_{k\to+\infty} 1. \tag{106}$$

A.2.2 The case w < 2

In this case the saddle-point (101) is outside the range of integration. The integral is dominated by $0 \le x \ll 1$, where

$$S(k, z, x) = -z - 2k \ln(-z) + \frac{k^w}{z} - \left(k + \frac{k^w}{z}\right) x + O(x^2), \tag{107}$$

so that

$$\int_0^{+\infty} \mathrm{d}x \mathrm{e}^{S(k,z,x)} \sim \frac{\exp\left[-z - 2k \ln(-z) + \frac{k^w}{z}\right]}{k + \frac{k^w}{z}} \cdot \tag{108}$$

We just have to look for the saddle-point of $\Sigma(k, z) = -z - 2k \ln(-z) + k^w/z$. The stationarity of Σ with respect to z reads

$$-1 - \frac{2k}{z} - \frac{k^w}{z^2} = 0. \tag{109}$$

There are two candidates as a saddle-point:

$$z_{sp}^{\pm} = -k \pm \sqrt{k^2 - k^w}. \tag{110}$$

It is easy to see that, when $k \to +\infty$, $\Sigma''(z_{sp}^+) < 0$ whereas $\Sigma''(z_{sp}^-) > 0$. So on a contour of the shape of Figure 3 the right saddle-point is z_{sp}^-. We have $\Sigma(k, z_{sp}^-) = -2k \ln k + O(k)$, so that

$$\int_C \mathrm{d}z \int_0^{+\infty} \mathrm{d}x \mathrm{e}^{S(k,z,x)} = O\left[\mathrm{e}^{-2k \ln k + O(k)}\right]. \tag{111}$$

Remembering (98), it follows that

$$g(k, w) = O\left[\mathrm{e}^{k(w-2)\ln k + O(k)}\right]. \tag{112}$$

So $g(k, w)$ goes to 0 when $k \to +\infty$.

Appendix B: The bipartite case

B.1 Computation of the averaged replicated partition function

We have two sets of N points each. We introduce the occupation numbers $n_{ij} = 0$ or 1, which are constrained by

$$\forall i \in 1, \dots N, \quad \sum_{j=1}^{N} n_{ij} = \sum_{j=1}^{N} n_{ji} = 1. \tag{113}$$

The length of the matching associated to a choice of the n_{ij} is $L(\{n_{ij}\}) = \sum_{i,j} n_{ij} l_{ij}$. The partition function (57) reads

$$Z = \sum_{n_{ij}=0,1} \int_0^{2\pi} \frac{\mathrm{d}\lambda_1}{2\pi} \cdots \int_0^{2\pi} \frac{\mathrm{d}\lambda_N}{2\pi} \int_0^{2\pi} \frac{\mathrm{d}\mu_1}{2\pi} \cdots \int_0^{2\pi} \frac{\mathrm{d}\mu_N}{2\pi}$$
$$\times \prod_i \exp\left(\mathrm{i}\lambda_i \left(1 - \sum_{j=1}^{N} n_{ij}\right)\right)$$
$$\times \prod_i \exp\left(\mathrm{i}\mu_i \left(1 - \sum_{j=1}^{N} n_{ji}\right)\right) \exp\left(-\frac{N}{2}\beta \sum_{i,j} n_{ij} l_{ij}\right), \tag{114}$$

where we enforced (113) using an integral representation of the Kronecker symbol $\delta(p) = \int_0^{2\pi} \mathrm{d}\lambda/(2\pi) \mathrm{e}^{\mathrm{i}p\lambda}$. It follows that

$$Z^n = \int \mathrm{d}[\lambda]\mathrm{d}[\mu] \exp\left(\mathrm{i} \sum_a \sum_i \lambda_i^a + \mu_i^a\right)$$
$$\times \prod_{i,j} \left(1 + \sum_\alpha \exp\left(-\frac{N}{2} p(\alpha)\beta l_{ij} - \mathrm{i} \sum_{a\in\alpha} (\lambda_i^a + \mu_j^a)\right)\right) \tag{115}$$

where $\mathrm{d}[\lambda]$ is a shorthand notation for $\prod_{a=1}^{n} \prod_{i=1}^{N} \mathrm{d}\lambda_i^a/(2\pi)$ (see the beginning of Sect. 2 for any precision on the other notations).

As we are interested in the subleading contribution to the free energy, it is necessary, when averaging over the disorder, to specify the distribution of the lengths we are considering:

$$\overline{\exp\left(-\frac{N}{2} p\beta l\right)} = \frac{2}{N} g_p - 4\frac{\mu}{N^2} g_p^2 + o\left(\frac{1}{N^2}\right), \tag{116}$$

where $\mu = 1$ in the case of the exponential distribution, 0 in the case of the flat distribution. Thus

$$\overline{Z^n} = \int \mathrm{d}[\lambda]\mathrm{d}[\mu] \exp\left(\mathrm{i} \sum_a \sum_i \lambda_i^a + \mu_i^a\right)$$
$$\times \prod_{i,j} \left(1 + 2\frac{T_{ij}}{N} - 4\mu \frac{R_{ij}}{N^2} + o\left(\frac{1}{N^2}\right)\right), \tag{117}$$

where

$$T_{ij} = \sum_\alpha g_\alpha \exp\left(-\mathrm{i} \sum_{a\in\alpha} \lambda_i^a + \mu_j^a\right),$$
$$R_{ij} = \sum_\alpha g_\alpha^2 \exp\left(-\mathrm{i} \sum_{a\in\alpha} \lambda_i^a + \mu_j^a\right). \tag{118}$$

Now we write

$$\prod_{i,j}\left(1+2\frac{1}{N}T_{ij}-4\frac{\mu}{N^2}R_{ij}\right)=\exp\left(\frac{2}{N}\sum_{i,j}T_{ij}\right)$$
$$\times\exp\left(-\frac{2}{N^2}\sum_{i,j}T_{ij}^2\right)\exp\left(-4\frac{\mu}{N^2}\sum_{i,j}R_{ij}\right)\dots \quad (119)$$

Plugged into (117), this gives

$$\overline{Z^n}=\int \mathrm{d}[\lambda]\mathrm{d}[\mu]\exp\left(\mathrm{i}\sum_a\sum_i \lambda_i^a+\mu_i^a\right)$$
$$\times\exp\left(\frac{1}{2N}\sum_\alpha g_\alpha\left(x_\alpha^2+y_\alpha^2\right)\right)$$
$$\times\exp\left(-\frac{1}{2N^2}\sum_{\alpha,\gamma}{}' g_\alpha g_\gamma\left(x_{\alpha\cup\gamma}^2+y_{\alpha\cup\gamma}^2\right)\right)$$
$$\times\exp\left(-\frac{\mu}{N^2}\sum_\alpha g_\alpha^2\left(x_\alpha^2+y_\alpha^2\right)\right), \quad (120)$$

where we set

$$\sum_i\exp\left(-\mathrm{i}\sum_{a\in\alpha}\lambda_i^a\right)=(x_\alpha+\mathrm{i}y_\alpha)/2, \quad (121)$$
$$\sum_i\exp\left(-\mathrm{i}\sum_{a\in\alpha}\mu_i^a\right)=(x_\alpha-\mathrm{i}y_\alpha)/2.$$

In (120), $\sum_{i,j}T_{ij}^2$ gives the only contribution $\sum'_{\alpha,\gamma}$ because the other terms vanish when integrated (remember that for p integer, $\int \mathrm{d}\lambda e^{\mathrm{i}p\lambda}=0$ unless $p=0$).

Using well known properties of Gaussian integrals, we can write

$$\overline{Z^n}=\int \mathrm{d}[\lambda]\mathrm{d}[\mu]\exp\left(\mathrm{i}\sum_a\sum_i \lambda_i^a+\mu_i^a\right)$$
$$\times\int\prod_\alpha \mathrm{d}X_\alpha\mathrm{d}Y_\alpha\frac{N}{2\pi g_\alpha}\exp\left(-\frac{N}{2}\sum_\alpha\frac{X_\alpha^2+Y_\alpha^2}{g_\alpha}\right)$$
$$\times\exp\left(\sum_\alpha X_\alpha x_\alpha+Y_\alpha y_\alpha\right)$$
$$\times\exp\left(-\frac{1}{2}\sum_{\alpha,\gamma}{}'\frac{g_\alpha g_\gamma}{g_{\alpha\cup\gamma}^2}\left(X_{\alpha\cup\gamma}^2+Y_{\alpha\cup\gamma}^2\right)\right)$$
$$\times\exp\left(-\mu\sum_\alpha\left(X_\alpha^2+Y_\alpha^2\right)\right). \quad (122)$$

Eventually, expressing x_α and y_α as functions of the λ_i^a and μ_i^a one gets

$$\overline{Z^n}=\int\prod_\alpha \mathrm{d}X_\alpha\mathrm{d}Y_\alpha\frac{N}{2\pi g_\alpha}\exp\left(-\frac{N}{2}\sum_\alpha\frac{X_\alpha^2+Y_\alpha^2}{g_\alpha}\right)$$
$$\times z[X_\alpha-\mathrm{i}Y_\alpha]^N z[X_\alpha+\mathrm{i}Y_\alpha]^N$$
$$\times\exp\left(-\frac{1}{2}\sum_{\alpha,\gamma}{}'\frac{g_\alpha g_\gamma}{g_{\alpha\cup\gamma}^2}\left(X_{\alpha\cup\gamma}^2+Y_{\alpha\cup\gamma}^2\right)\right)$$
$$\times\exp\left(-\mu\sum_\alpha\left(X_\alpha^2+Y_\alpha^2\right)\right). \quad (123)$$

B.2 Computation of ΔF_2^1 (64)

Let us recall a result of [2]

$$Q_p^{sp}=\frac{2}{\beta p}\int_{-\infty}^{+\infty}\mathrm{d}l\frac{e^{lp}}{(p-1)!}e^{-G(l)}. \quad (124)$$

Using (62) we see that, when $n\to 0$ and $\beta\to+\infty$,

$$\frac{1}{nN\beta}\sum_\alpha (Q_\alpha^{sp})^2=\frac{1}{N\beta}\sum_{p=1}^{+\infty}\frac{(-1)^{p-1}}{p}(Q_p^{sp})^2$$
$$=\frac{2}{N\beta^2}\int_{-\infty}^{+\infty}\mathrm{d}l\sum_p Q_p^{sp}\frac{(-1)^{p-1}}{pp!}e^{pl}e^{-G(l)}$$
$$=\frac{2}{N\beta^2}\int_{-\infty}^{+\infty}\mathrm{d}lG(l)\int_l^{+\infty}\mathrm{d}te^{-G(t)} \quad \text{by (11)}$$
$$=\frac{1}{2N}\int_0^{+\infty}\mathrm{d}u\ln(1+u)\ln(1+1/u) \quad \text{by (39)}$$
$$=\frac{1}{N}\zeta(3). \quad (125)$$

B.3 Computation of the volume of the hypersurface of saddle-points (72)

The volume is $\int_0^{2\pi}\mathrm{d}\theta_1\dots\int_0^{2\pi}\mathrm{d}\theta_n\sqrt{\det g}$ where

$$g_{ab}=\sum_\alpha\left[\frac{\partial U_\alpha^{sp}}{\partial\theta_a}\frac{\partial U_\alpha^{sp}}{\partial\theta_b}+\frac{\partial V_\alpha^{sp}}{\partial\theta_a}\frac{\partial V_\alpha^{sp}}{\partial\theta_b}\right]. \quad (126)$$

g has a very simple structure: all diagonal elements are equal to g_0, all non diagonal elements to g_1, with

$$g_0=\sum_{p=1}^{+\infty}\left(U_p^{sp}\right)^2C_{n-1}^{p-1}\sim_{n\to 0}\sum_{p=1}^{+\infty}(-1)^{p-1}\frac{\left(Q_p^{sp}\right)^2}{g_p}, \quad (127)$$
$$g_1=\sum_{p=2}^{+\infty}\left(U_p^{sp}\right)^2C_{n-2}^{p-2}\sim_{n\to 0}\beta\sum_{p=1}^{+\infty}(-1)^p p^2\left(Q_p^{sp}\right)^2+g_0. \quad (128)$$

Using (62) and (124), one gets

$$g_0 = \beta \sum_{p=1}^{+\infty} (-1)^{p-1} p Q_p^{sp} \frac{2}{\beta p} \int_{-\infty}^{+\infty} \mathrm{d}l \frac{\mathrm{e}^{lp}}{(p-1)!} \mathrm{e}^{-G(l)}$$
$$= 2 \int_{-\infty}^{+\infty} \mathrm{d}l G'(l) \mathrm{e}^{-G(l)} = 2 \quad (129)$$

and

$$g_1 - g_0 = 2 \sum_{p=1}^{+\infty} (-1)^{p} p Q_p^{sp} \int_{-\infty}^{+\infty} \mathrm{d}l \frac{\mathrm{e}^{lp}}{(p-1)!} \mathrm{e}^{-G(l)}$$
$$= -2 \int_{-\infty}^{+\infty} \mathrm{d}l G''(l) \mathrm{e}^{-G(l)} = -\frac{2}{\beta} \cdot \quad (130)$$

The computation of $\det g$ gives

$$\det = [g_0(n) - g_1(n)]^{n-1} [g_0(n) + (n-1) g_1(n)], \quad (131)$$

so that

$$\sqrt{\det g} = 1 + \frac{n}{2} \left[\frac{g_1}{g_0 - g_1} + \ln(g_0 - g_1) \right] + o(n) \quad (132)$$

(here again one must be careful when deriving this result that there might *a priori* be some contribution of $\mathrm{d}g_0/\mathrm{d}n$ or $\mathrm{d}g_1/\mathrm{d}n$).

Thus the contribution to the free energy is

$$\Delta F^3 = -\frac{1}{2N\beta} \left[\frac{g_1}{g_0 - g_1} + \ln(g_0 - g_1) \right]$$
$$=_{\beta \to +\infty} -\frac{1}{2N} \cdot \quad (133)$$

Mézard and Parisi [8] had expressions (127) and (128) with X_p^{sp} instead of U_p^{sp}, which made them find a wrong $\Delta F^3 = -\pi^2/(24N)$.

References

1. O.C. Martin, R. Monasson, R. Zecchina, Theor. Comp. Sci. **265**, 3 (2001)
2. M. Mézard, G. Parisi, J. Phys. Lett. **46**, L-771 (1985)
3. M. Mézard, G. Parisi, Europhys. Lett. **2**, 913 (1986)
4. R. Brunetti, W. Krauth, M. Mézard, G. Parisi, Europhys. Lett. **14**, 295 (1991)
5. J. Houdayer, J.H. Boutet de Monvel, O.C. Martin, Eur. Phys. J. B **6**, 383 (1998)
6. G. Parisi, M. Ratiéville, Eur. Phys. J. B **22**, 229 (2001)
7. D.J. Aldous, `math.PR/0010063`, and Random Struct. Algorithms **18**, 381 (2001)
8. M. Mézard, G. Parisi, J. Phys. France **48**, 1451 (1987)
9. G. Parisi, *A Conjecture on random bipartite matching*, `cond-mat/9801176`
10. J. Houdayer, Ph.D. thesis, University Paris XI, France (1999)
11. V. Dotsenko, J. Phys. A **33**, 2015 (2000)
12. J. Vannimenus, M. Mézard, J. Phys. Lett. **45**, L-1145 (1984)
13. J.R.L. de Almeida, D. Thouless, J. Phys. A **11**, 983 (1978)
14. G. Parisi, *Spin glasses and optimization problems without replicas. Les Houches, Session XLVI*, 1986 — Chance and Matter, edited by J. Souletie *et al.* (Elsevier Science Publishers B.V., 1987), p. 525
15. M. Mézard, *Spin glasses and optimization. Invited paper at the Heidelberg colloquium on glassy dynamics and optimization*, Lecture Notes in Physics **275** (Springer Verlag 1987), p. 354
16. W. Krauth, M. Mézard, Europhys. Lett. **8**, 213 (1989)
17. M. Mézard, G. Parisi, J. Phys. France **49**, 2019 (1988)
18. N.J. Cerf, J.H. Boutet de Monvel, O.C. Martin, Phys. Rev. Lett. **79**, 167 (1997)
19. P. Grassberger, H. Freund, Z. Oper. Res. **34**, 239 (1990)
20. C. Bender, S. Orszag, *Advanced Mathematical Methods for Scientists and Engineers*, International Series in Pure and Applied Mathematics (McGraw-Hill, Inc., New York, 1978)