7. An Assignment Problem

7.1 Introduction

Given positive numbers $c(i,j)$, $i,j \le N$, the assignment problem is to find

$$\min_\sigma \sum_{i \le N} c(i,\sigma(i)) \,, \tag{7.1}$$

where σ ranges over all permutations of $\{1,\ldots,N\}$. In words, if $c(i,j)$ represents the cost of assigning job j to worker i, we want to minimize the total cost when exactly one job is assigned to each worker.

We shall be interested in the random version of the problem, where the numbers $c(i,j)$ are independent and uniformly distributed over $[0,1]$.

Mézard and Parisi [103], [104] studied (7.1) by introducing a suitable Hamiltonian, and conjectured that

$$\lim_{N\to\infty} \mathsf{E} \min_\sigma \sum_{i\le N} c(i,\sigma(i)) = \frac{\pi^2}{6} \,. \tag{7.2}$$

This was proved by D. Aldous [2]. Aldous takes advantage of a feature of the present model, that makes it rather special among the various models we studied: the existence of a "limiting object" (which he discovered [1]).

In a related direction, G. Parisi conjectured the following remarkable identity. If the r.v.s $c(i,j)$ are independent exponential i.e. they satisfy $\mathsf{P}(c(i,j) \ge x) = e^{-x}$ for $x \ge 0$, then we have

$$\mathsf{E} \min_\sigma \sum_{i\le N} c(i,\sigma(i)) = 1 + \frac{1}{2^2} + \cdots + \frac{1}{N^2} \,. \tag{7.3}$$

The link with (7.2) is that it can be shown that if the r.v.s $c(i,j)$ are i.i.d., and their common distribution has a density f on $\mathbb{R}^+$ with respect to Lebesgue measure, then if f is continuous in a neighborhood of 0, the limit in (7.2) depends only on $f(0)$. (The intuition for this is simply that all the numbers $c(i,\sigma(i))$ relevant in the computation of the minimum in (7.2) should be very small for large N, so that only the part of the distribution of $c(i,j)$ close to 0 matters.) Thus it makes no difference to assume that $c(i,j)$ is uniform over $[0,1]$ or is exponential of mean 1.

M. Talagrand, *Mean Field Models for Spin Glasses*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics 54, DOI 10.1007/978-3-642-15202-3_7,

Vast generalizations of Parisi's conjecture have been recently proved [109], [96]. Yet the disordered system introduced by Mézard and Parisi remains of interest. This model is obviously akin to the other models we consider; yet it is rather different. In the author's opinion, this model demonstrates well the far-reaching nature of the ideas underlying the theory of mean field models for spin glasses.

It is a great technical challenge to prove rigorously anything at all concerning the original model of Mézard and Parisi. This challenge has yet to be met. We will consider a slightly different model, that turns out to be easier, but still of considerable interest. In this model, we consider two integers M, N, $M \geq N$. We consider independent r.v.s $(c(i,j))_{i \leq N, j \leq M}$ that are uniform over $[0,1]$. The configuration space is the set $\Sigma_{N,M}$ of all one-to-one maps σ from $\{1, \ldots, N\}$ to $\{1, \ldots, M\}$. On this space we consider the Hamiltonian

$$H_{N,M}(\sigma) = \beta N \sum_{i \leq N} c(i, \sigma(i)) \ , \tag{7.4}$$

where β is a parameter. The reader observes that there is no minus sign in this formula, that is, the Boltzmann factor is

$$\exp\left(-\beta N \sum_{i \leq N} c(i, \sigma(i))\right) .$$

Given a number $\alpha > 0$, we will study the system for $N \to \infty$, $M = \lfloor N(1+\alpha) \rfloor$, and our results will hold for $\beta \leq \beta(\alpha)$, where, unfortunately, $\lim_{\alpha \to 0} \beta(\alpha) = 0$. The original model of Mézard and Parisi is the case $M = N$, i.e. $\alpha = 0$. A step towards understanding this model would be the following.

Research Problem 7.1.1. (Level 2) Extend the results of the present chapter to the case $\beta \leq \beta_0$ where β_0 is independent of α.

Even in the domain $\beta \leq \beta(\alpha)$ our results are in a sense weaker than those of the previous chapters. We do not study the model for given large values of N and M, but only in the limit $N \to \infty$ and $M/N \to \alpha$, and we do not obtain a rate for several of the convergence results.

One of the challenges of the present situation is that it is not obvious how to formulate the correct questions. We expect (under our condition that β is small) that "the spins at two different sites are nearly independent". Here this should mean that when $i_1 \neq i_2$, under Gibbs' measure the variables $\sigma \mapsto \sigma(i_1)$ and $\sigma \mapsto \sigma(i_2)$ are nearly independent. But how could one quantify this phenomenon in a way suitable for a proof by induction?

We consider the partition function

$$Z_{N,M} = \sum_{\sigma} \exp(-H_{N,M}(\sigma)) \ , \tag{7.5}$$

where the summation is over all possible values of σ in $\Sigma_{N,M}$. Throughout the chapter we write

$$a(i,j) = \exp(-\beta N c(i,j)) , \tag{7.6}$$

so that

$$Z_{N,M} = \sum_{\sigma} \prod_{i \le N} a(i, \sigma(i)) .$$

The cavity method will require removing elements from $\{1, \dots, N\}$ and $\{1, \dots, M\}$. Given a set $A \subset \{1, \dots, N\}$ and a set $B \subset \{1, \dots, M\}$ such that $N - \operatorname{card} A \le M - \operatorname{card} B$, we write

$$Z_{N,M}(A; B) = \sum_{\sigma} \prod a(i, \sigma(i)) .$$

The product is taken over $i \in \{1, \dots, N\} \backslash A$ and the sum is taken over the one-to-one maps σ from $\{1, \dots, N\} \backslash A$ to $\{1, \dots, M\} \backslash B$. Thus $Z_{N,M} = Z_{N,M}(\emptyset; \emptyset)$. When $A = \{i_1, i_2, \dots\}$ and $B = \{j_1, j_2, \dots\}$ we write

$$Z_{N,M}(A, B) = Z_{N,M}(i_1, i_2, \dots; j_1, j_2, \dots) .$$

Rather than working directly with Gibbs' measure, we will prove that

$$\frac{Z_{N,M}(i;j)}{Z_{N,M}} \simeq \frac{Z_{N,M}(\emptyset;j)}{Z_{N,M}} \frac{Z_{N,M}(i;\emptyset)}{Z_{N,M}} . \tag{7.7}$$

It should be obvious that this is a very strong property, and that it deals with independence. One can also get convinced that it deals with Gibbs' measure by observing that

$$G(\{\sigma(i) = j\}) = a(i,j) \frac{Z_{N,M}(i,j)}{Z_{N,M}} .$$

We consider the quantities

$$u_{N,M}(j) = \frac{Z_{N,M}(\emptyset;j)}{Z_{N,M}} ; \; w_{N,M}(i) = \frac{Z_{N,M}(i;\emptyset)}{Z_{N,M}} . \tag{7.8}$$

These quantities occur in the right-hand side of (7.7). The number $u_{N,M}(j)$ is the Gibbs probability that j does not belong to the image of $\{1, \dots, N\}$ under the map σ. In particular we have $0 \le u_{N,M}(j) \le 1$. (On the other hand we only know that $w_{N,M}(i) > 0$.)

Having understood that these quantities are important, we would like to know something about the family $(u_{N,M}(j))_{j \le M}$ (or $(w_{N,M}(i))_{i \le N}$). An optimistic thought is that this family looks like an i.i.d. sequence drawn out of a certain distribution, that we would like to describe, probably as a fixed point of a certain operator. Analyzing the problem, it is not very difficult to guess what the operator should be; the unpleasant surprise is that it does not seem obvious that this operator has a fixed point, and this contributes significantly to the difficulty of the problem. In order to state our main result, let us describe this operator. Of course, the motivation behind this definition will become clear only gradually.

Consider a standard Poisson point process on $\mathbb{R}^+$ (that is, its intensity measure is Lebesgue's measure) and denote by $(\xi_i)_{i\geq 1}$ an increasing enumeration of the points it produces. Consider a probability measure η on $\mathbb{R}^+$, and i.i.d. r.v.s $(Y_i)_{i\geq 1}$ distributed according to η, which are independent of the r.v.s ξ_i. We define

$$A(\eta) = \mathcal{L}\left(\frac{1}{\sum_{i\geq 1} Y_i \exp\left(-\beta\xi_i/(1+\alpha)\right)}\right) \tag{7.9}$$

$$B(\eta) = \mathcal{L}\left(\frac{1}{1+\sum_{i\geq 1} Y_i \exp(-\beta\xi_i)}\right), \tag{7.10}$$

where of course $\mathcal{L}(X)$ is the law of the r.v. X. The dependence on β and α is kept implicit.

Theorem 7.1.2. *Given $\alpha > 0$, there exists $\beta(\alpha) > 0$ such that for $\beta \leq \beta(\alpha)$ there exists a unique pair μ, ν where μ is a probability measure on $[0,1]$ and ν is a probability measure on $\mathbb{R}^+$ such that*

$$\int x \mathrm{d}\mu(x) = \frac{\alpha}{1+\alpha} \ ; \ \mu = B(\nu) \ ; \ \nu = A(\mu) \ . \tag{7.11}$$

Moreover if $M = \lfloor N(1+\alpha) \rfloor$, we have

$$\mu = \lim_{N\to\infty} \mathcal{L}(u_{N,M}(M)) \ ; \ \nu = \lim_{N\to\infty} \mathcal{L}(w_{N,M}(N)) \ . \tag{7.12}$$

Research Problem 7.1.3. (Level 2) Find a direct proof of the existence of the pair (μ,ν) as in (7.11).

One intrinsic difficulty is that there exists such a pair for each value of α (not too small); so one cannot expect that the operator $B \circ A$ is a contraction for a certain distance. The way we will prove (7.11) is by showing that a cluster point of the sequence $(\mathcal{L}(u_{N,M}(M)), \mathcal{L}(w_{N,M}(N)))$ is a solution of these equations.

While it is not entirely obvious what are the relevant questions one should ask about the system, the following shows that the objects of Theorem 7.1.2 are of central importance.

Theorem 7.1.4. *Given α, for $\beta \leq \beta(\alpha)$ we have*

$$\lim_{N\to\infty} \frac{1}{N} \mathsf{E} \log Z_{N,M} = -(1+\alpha)\int \log x \,\mathrm{d}\mu(x) - \int \log x \,\mathrm{d}\nu(x) \ . \tag{7.13}$$

7.2 Overview of the Proof

In this section we try to describe the overall strategy. The following fundamental identities are proved in Lemma 7.3.4 below

$$u_{N,M}(M) = \frac{1}{1 + \sum_{k \le N} a(k,M)\, w_{N,M-1}(k)} \tag{7.14}$$

$$w_{N,M}(N) = \frac{1}{\sum_{\ell \le M} a(N,\ell)\, u_{N-1,M}(\ell)} \,. \tag{7.15}$$

Observe that in the right-hand side of (7.14) the r.v.s $a(k,M)$ are independent of the numbers $w_{N,M-1}(k)$, and similarly in (7.15). We shall prove that

$$w_{N,M}(k) \simeq w_{N,M-1}(k) \simeq w_{N,M-2}(k) \,. \tag{7.16}$$

This fact is not easy. It is intimately connected to equation (7.7), and is rigorously established in Theorem 7.4.7 below.

Once we have (7.16) we see from (7.14) that

$$u_{N,M}(M) \simeq \frac{1}{1 + \sum_{k \le N} a(k,M)\, w_{N,M-2}(k)} \,, \tag{7.17}$$

and by symmetry between M and $M-1$ that

$$u_{N,M}(M-1) \simeq \frac{1}{1 + \sum_{k \le N} a(k,M-1)\, w_{N,M-2}(k)} \,. \tag{7.18}$$

As a consequence, given the numbers $w_{N,M-2}(k)$, the r.v.s $u_{N,M}(M)$ and $u_{N,M}(M-1)$ are nearly independent. Their common law depends only on the empirical measure

$$\frac{1}{N} \sum_{i \le N} \delta_{w_{N,M-2}(i)} \,,$$

which, by (7.16), is nearly

$$\nu_N = \frac{1}{N} \sum_{i \le N} \delta_{w_{N,M}(i)} \,. \tag{7.19}$$

We consider an independent sequence of r.v.s $(X_k)_{k \ge 1}$ uniformly distributed on $[0,1]$, independent of all the other sources of randomness, and we set

$$a(k) = \exp(-\beta N X_k) \,. \tag{7.20}$$

The reason this sequence is of fundamental importance for the present model is that, given j, the sequence $(a(k,j))_k$ of r.v.s has the same distribution as the sequence $(a(k))_k$, and, given i, this is also the case of the sequence $(a(i,k))_k$.

Consider the random measure $\overline{\mu}_N$ on $[0,1]$ given by

$$\overline{\mu}_N = \mathcal{L}_a\left(\frac{1}{1+\sum_{k\le N} a(k) w_{N,M}(k)}\right) ,$$

where $\mathcal{L}_a$ denotes the law in the randomness of the variables $a(k)$, when all the other sources of randomness are fixed.

Thus, given the numbers $w_{N,M}(k)$, the r.v.s $u_{N,M}(M)$ and $u_{N,M}(M-1)$ are nearly independent with common law $\overline{\mu}_N$. By symmetry this is true for each pair of r.v.s $u_{N,M}(j)$ and $u_{N,M}(k)$.

Therefore we expect that the empirical measure

$$\mu_N = \frac{1}{M} \sum_{j\le M} \delta_{u_{N,M}(j)}$$

is nearly $\overline{\mu}_N$.

Since $\overline{\mu}_N$ is a continuous function of ν_N, it follows that if ν_N is concentrated (in the sense that it is nearly non-random), then such is the case of $\overline{\mu}_N$, that is nearly concentrated around its mean μ'_N, and therefore μ_N itself is concentrated around μ'_N.

We can argue similarly that if μ_N is concentrated around μ'_N, then ν_N must be concentrated around a certain measure ν'_N that can be calculated from μ_N. The hard part of the proof is to get quantitative estimates showing that if β is sufficiently small, then these cross-referential statements can be combined to show that both μ_N and ν_N are concentrated around μ'_N and ν'_N respectively. Now, the way μ'_N is obtained from ν'_N means in the limit that $\mu'_N \simeq B(\nu'_N)$. Similarly, $\nu'_N \simeq A(\mu'_N)$. Also, $\mu'_N = \mathcal{L}(u_{N,M}(M))$ and $\nu'_M = \mathcal{L}(w_{N,M}(N))$, so $\mu = \lim_N \mathcal{L}(u_{N,M}(M))$ and $\nu = \lim_N \mathcal{L}(w_{N,M}(N))$ satisfy $\mu = B(\nu)$ and $\nu = A(\mu)$.

7.3 The Cavity Method

We first collect some simple facts.

Lemma 7.3.1. *If $i \notin A$, we have*

$$Z_{N,M}(A;B) = \sum_{\ell\notin B} a(i,\ell)\, Z_{N,M}(A\cup\{i\}\,;\ B\cup\{\ell\}) \ . \tag{7.21}$$

If $j \notin B$, we have

$$Z_{N,M}(A;B) = Z_{N,M}(A;B\cup\{j\}) + \sum_{k\notin A} a(k,j)\, Z_{N,M}(A\cup\{k\}\,;\ B\cup\{j\}) \ . \tag{7.22}$$

Proof. One replaces each occurrence of $Z_{N,M}(\cdot;\cdot)$ by its value and one checks that the same terms occur in the left-hand and right-hand sides. $\square$

The following deserves no proof.

Lemma 7.3.2. *If $M \notin B$, we have*

$$Z_{N,M}(A; B \cup \{M\}) = Z_{N,M-1}(A; B) . \tag{7.23}$$

If $N \notin A$, we have

$$Z_{N,M}(A \cup \{N\}; B) = Z_{N-1,M}(A; B) . \tag{7.24}$$

In (7.24), and in similar situations below, we make the convention that $Z_{N-1,M}(\cdot;\cdot)$ is considered for a parameter β' such that $\beta'(N-1) = \beta N$.

The following is also obvious from the definitions, yet it is fundamental.

Lemma 7.3.3. *We have*

$$\sum_{\ell \le M} Z_{N,M}(\emptyset; \ell) = (M-N) Z_{N,M} \tag{7.25}$$

and thus

$$\sum_{\ell \le M} u_{N,M}(\ell) = M - N . \tag{7.26}$$

To prove (7.26) we can also observe that $u_{N,M}(\ell)$ is the Gibbs probability that ℓ does not belong to the image under σ of $\{1, \cdots, N\}$, so that the left-hand side of (7.26) is the expected number of integers that do not belong to this image, i.e. $M - N$. In particular (7.26) implies by symmetry between the values of ℓ that $\mathsf{E} u_{N,M}(M) = (M-N)/M \simeq \alpha/(1+\alpha)$, so that any cluster point μ of the sequence $\mathcal{L}(u_{N,M}(M))$ satisfies $\int x \mathrm{d}\mu(x) = \alpha/(1+\alpha)$.

Lemma 7.3.4. *We have*

$$u_{N,M}(M) = \frac{Z_{N,M-1}}{Z_{N,M}} = \frac{1}{1 + \sum_{k \le N} a(k,M)\, w_{N,M-1}(k)} \tag{7.27}$$

$$w_{N,M}(N) = \frac{Z_{N-1,M}}{Z_{N,M}} = \frac{1}{\sum_{\ell \le M} a(N,\ell)\, u_{N-1,M}(\ell)} . \tag{7.28}$$

Proof. We use (7.22) with $A = B = \emptyset$ and $j = M$ to obtain

$$Z_{N,M} = Z_{N,M}(\emptyset; M) + \sum_{k \le N} a(k,M)\, Z_{N,M}(k; M) .$$

Using (7.23) with $A = \emptyset$ or $A = \{k\}$ and $B = \emptyset$ we get

$$\begin{aligned} Z_{N,M} &= Z_{N,M-1} + \sum_{k \le N} a(k,M)\, Z_{N,M-1}(k;\emptyset) \\ &= Z_{N,M-1}\left(1 + \sum_{k\le N} a(k,M)\, w_{N,M-1}(k)\right) . \end{aligned} \tag{7.29}$$

This proves (7.27). The proof of (7.28) is similar, using now (7.21) and (7.24). □

It will be essential to consider the following quantity, where $i \le N$:

$$L_{N,M}(i) = \frac{Z_{N,M}\, Z_{N,M-1}(i;\emptyset) - Z_{N,M}(i;\emptyset)\, Z_{N,M-1}}{Z_{N,M}^2} . \tag{7.30}$$

The idea is that (7.7) used for $j = M$ implies that $\mathsf{E} L_{N,M}(i)^2$ is small. (This expectation does not depend on i.) Conversely, if $\mathsf{E} L_{N,M}(i)^2$ is small this implies (7.7) for $j = M$ and hence for all values of j by symmetry.

We will also use the quantity

$$R_{N,M}(j) = \frac{Z_{N,M}\, Z_{N,M-1}(\emptyset;j) - Z_{N,M}(\emptyset;j)\, Z_{N,M-1}}{Z_{N,M}^2} . \tag{7.31}$$

It is good to notice that $|R_{N,M}(j)| \le 2$. This follows from (7.23) and the fact that the quantity $Z_{N,M}(A,B)$ decreases as B increases.

The reason for introducing the quantity $R_{N,M}(j)$ is that it occurs naturally when one tries to express $L_{M,N}(i)$ as a function of a smaller system (as the next lemma shows).

Lemma 7.3.5. *We have*

$$L_{N,M}(N) = -\frac{\sum_{\ell\le M-1} a(N,\ell)\, R_{N-1,M}(\ell) - a(N,M)\, u_{N-1,M}(M)^2}{\left(\sum_{\ell\le M} a(N,\ell)\, u_{N-1,M}(\ell)\right)^2} \tag{7.32}$$

$$R_{N,M}(M-1) = -\frac{\sum_{k\le N} a(k,M)\, L_{N,M-1}(k)}{\left(1+\sum_{k\le N} a(k,M)\, w_{N,M-1}(k)\right)^2} . \tag{7.33}$$

Proof. Using the definition (7.31) of $R_{N,M}(j)$ with $j = M-1$, we have

$$R_{N,M}(M-1) = \frac{Z_{N,M}\, Z_{N,M-1}(\emptyset;M-1) - Z_{N,M}(\emptyset;M-1)\, Z_{N,M-1}}{Z_{N,M}^2} . \tag{7.34}$$

As in (7.29), but using now (7.22) with $B = \{M-1\}$ and $j = M$ we obtain:

$$\begin{aligned} Z_{N,M}(\emptyset;M-1) &= Z_{N,M-1}(\emptyset;M-1) \\ &\quad + \sum_{k\le N} a(k,M)\, Z_{N,M-1}(k;M-1) . \end{aligned} \tag{7.35}$$

Using this and (7.29) in the numerator of (7.34), and (7.29) in the denominator, and gathering the terms yields (7.33). The proof of (7.32) is similar. □

We end this section by a technical but essential fact.

Lemma 7.3.6. *We have*

$$\sum_{j\le M-1} R_{N,M}(j) = -u_{N,M}(M) + u_{N,M}(M)^2 \,. \tag{7.36}$$

Proof. From (7.25) we have

$$\sum_{j\le M-1} Z_{N,M}(\emptyset;j) = (M-N)\,Z_{N,M} - Z_{N,M}(\emptyset;M) = (M-N)\,Z_{N,M} - Z_{N,M-1},$$

and changing M into $M-1$ in (7.25) we get

$$\sum_{j\le M-1} Z_{N,M-1}(\emptyset,j) = (M-1-N)\,Z_{N,M-1} \,.$$

These two relations imply (7.36) in a straightforward manner. □

7.4 Decoupling

In this section, we prove (7.7) and, more precisely, the following.

Theorem 7.4.1. *Given $\alpha > 0$, there exists $\beta(\alpha) > 0$ such that if $\beta \le \beta(\alpha)$ and $M = \lfloor N(1+\alpha)\rfloor$, then for $\beta N \ge 1$*

$$\mathsf{E}\, L_{N,M}(N)^2 \le \frac{K(\alpha)}{N} \tag{7.37}$$

$$\mathsf{E}\, R_{N,M}(M-1)^2 \le \frac{K(\alpha)}{N} \,. \tag{7.38}$$

The method of proof consists of using Lemma 7.3.5 to relate $\mathsf{E}R_{N,M}(M-1)^2$ with $\mathsf{E}L_{N,M-1}(N)^2$ and $\mathsf{E}L_{N,M}(N)^2$ with $\mathsf{E}R_{N-1,M}(M-1)^2$, and to iterate these relations. In the right-hand sides of (7.32) and (7.33), we will first take expectation in the quantities $a(N,\ell)$ and $a(k,M)$, that are probabilistically independent of the other quantities (an essential fact). Our first task is to learn how to do this.

We recall the random sequence $a(k) = \exp(-\beta N X_k)$ of (7.20), where (X_k) are i.i.d., uniform over $[0,1]$, and independent of the other sources of randomness. The following lemma is obvious.

Lemma 7.4.2. *We have*

$$\mathsf{E}\, a(k)^p = \frac{1}{\beta p N}(1-\exp(-\beta p N)) \le \frac{1}{\beta p N} \,. \tag{7.39}$$

Lemma 7.4.3. *Consider numbers $(x_k)_{k\le N}$. Then we have*

$$\mathsf{E}\Big(\sum_{k\le N} a(k)\, x_k\Big)^2 \le \Big(\frac{1}{2\beta^2 N} + \frac{1}{2\beta N}\Big) \sum_{k\le N} x_k^2 \,. \tag{7.40}$$

Proof. Using (7.39) we have

$$\mathsf{E}\Bigl(\sum_{k\le N} a(k)\, x_k\Bigr)^2 = \sum_{k\le N} x_k^2\, \mathsf{E}\, a(k)^2 + \sum_{k\ne\ell} x_k\, x_\ell\, \mathsf{E}\, a(k)\, \mathsf{E}\, a(\ell)$$
$$\le \frac{1}{2\beta N}\sum_{k\le N} x_k^2 + \Bigl(\frac{1}{\beta N}\Bigr)^2 \sum_{k\ne\ell} |x_k|\,|x_\ell|\,.$$

Now, the Cauchy-Schwarz inequality implies:

$$\sum_{k\ne\ell} |x_k|\,|x_\ell| \le \frac{1}{2}\Bigl(\sum_{k\le N} |x_k|\Bigr)^2 \le \frac{N}{2}\sum_{k\le N} x_k^2\,.$$ □

Corollary 7.4.4. *If $\beta\le 1$ we have*

$$\mathsf{E}\, R_{N,M}(M-1)^2 \le \frac{1}{\beta^2}\, \mathsf{E}\, L_{N,M-1}(N)^2\,.$$

Proof. From (7.33) we have

$$R_{N,M}(M-1)^2 \le \Bigl(\sum_{k\le N} a(k,M)\, L_{N,M-1}(k)\Bigr)^2\,.$$

The sequence $(a(k,M))_{k\le N}$ has the same distribution as the sequence $(a(k))_{k\le N}$, so that taking expectation first in this sequence and using (7.40) we get, assuming without loss of generality that $\beta\le 1$,

$$\mathsf{E}\, R_{N,M}(M-1)^2 \le \frac{1}{\beta^2 N}\sum_{k\le N} \mathsf{E}\, L_{N,M-1}(k)^2 = \frac{1}{\beta^2}\, \mathsf{E}\, L_{N,M-1}(N)^2$$

by symmetry between the values of k. □

This is very crude because in (7.33) the denominator is not of order 1, but seems to be typically much larger. In order however to prove this, we need to know that a proportion of the numbers $(w_{N,M-1}(k))_{k\le M}$ are large. We will prove that this is indeed the case if $\beta\le\beta(\alpha)$, but we do not know it yet. To improve on the present approach it seems that we would need to have this information now. We could not overcome this technical difficulty, that seems related to Research Problem 7.1.1.

We next turn to the task of taking expectation in (7.32). The relation (7.26) is crucial here. Since $0\le u_{N,M}(\ell)\le 1$ and $M-N\simeq N\alpha$, this relation implies that at least a constant proportion of the numbers $(u(\ell))_{\ell\le M} = (u_{N,M}(\ell))_{\ell\le M}$ is not small. To understand what happens, consider an independent sequence X_ℓ uniformly distributed over $[0,1]$ and note that if we reorder the numbers $(NX_\ell)_{\ell\le M}$ by increasing order, they look like the sequence $(\xi_i/(1+\alpha))$ (where $(\xi_i)_{i\ge 1}$ is an enumeration of the points of a Poisson point process on $\mathbb{R}^+$). The sum $\sum_{\ell\le N} a(\ell)u(\ell)$ then looks like the sum $\sum_{\ell\le N}\exp(-\beta\xi_\ell/(1+\alpha))u(\sigma(\ell))$ where σ is a random permutation, and it is easy to get convinced that typically it cannot be too small. The precise technical result we need is as follows.

Proposition 7.4.5. *Consider numbers $0\le u(\ell), u'(\ell)\le 1$, for $\ell\le M$. Assume that $\sum_{\ell\le M} u(\ell)\ge 4$ and $\sum_{\ell\le M} u'(\ell)\ge 4$. Consider b with $Nb\le \sum_{\ell\le M} u(\ell)$ and $Nb\le\sum_{\ell\le M} u'(\ell)$. Then if $\beta N\ge 1$ and if $\beta\le b/40$, for any numbers $(y(\ell))_{\ell\le M}$ we have*

$$\mathsf{E}\frac{\left(\sum_{\ell\le M} a(\ell)y(\ell)\right)^2}{\left(\sum_{\ell\le M} a(\ell)u(\ell)\right)^2\left(\sum_{\ell\le M} a(\ell)u'(\ell)\right)^2}\le\frac{L\beta^2}{b^4}\left(\frac{1}{N}\sum_{\ell\le M} y(\ell)\right)^2 + \frac{L\beta^3}{b^6 N}\sum_{\ell\le M} y(\ell)^2\ , \tag{7.41}$$

where $a(\ell)=\exp(-\beta N X_\ell)$ and L denotes a universal constant.

As will be apparent later, an essential feature is that the second term of this bound has a coefficient β^3 (rather than β^2).

Corollary 7.4.6. *If $\beta\le\alpha/80$, $\beta N\ge 1$, $M\ge\lfloor N(1+\alpha)\rfloor$, $M\le 3N$, we have*

$$\mathsf{E}L_{N,M}(N)^2\le\frac{L\beta^3}{\alpha^6}\mathsf{E}R_{N-1,M}(M-1)^2+\frac{K(\alpha)}{N}\ . \tag{7.42}$$

Proof. For $\ell\le M$, let $u(\ell)=u_{N-1,M}(\ell)$, and $a(\ell)=a(N,\ell)$. For $\ell\le M-1$ let $y(\ell)=R_{N-1,M}(\ell)$, and let $y(M)=-u_{N-1,M}(M)^2$. By (7.32) we have

$$L_{N,M}(N)^2=\frac{\left(\sum_{\ell\le M} a(\ell)y(\ell)\right)^2}{\left(\sum_{\ell\le M} a(\ell)u(\ell)\right)^4}\ .$$

We check first that $\sum_{\ell\le M} u(\ell)\ge 4$. Then (7.26) implies

$$\sum_{\ell\le M} u(\ell)=M-(N-1)\ge\lfloor N(1+\alpha)\rfloor-N=\lfloor N\alpha\rfloor\ ,$$

and if $\beta\le\alpha/80$ and $\beta N\ge 1$, then $N\alpha\ge 80$ and this is certainly ≥ 4. Also

$$b:=\frac{1}{N}\sum_{\ell\le M} u(\ell)\ge\frac{\lfloor N\alpha\rfloor}{N}\ge\frac{\alpha}{2}$$

if $N\alpha\ge 2$ and in particular if $N\beta\ge 1$ and $\beta\le\alpha/80$. We then have $\beta\le b/40$. Taking expectation in the r.v.s $a(\ell)$, we can now use (7.41) with $u'(\ell)=u(\ell)$ to obtain

$$\mathsf{E}_a L_{N,M}(N)^2 \le \frac{L\beta^2}{\alpha^4}\left(\frac{1}{N}\sum_{\ell\le M} y(\ell)\right)^2 + \frac{L\beta^3}{\alpha^6 N}\sum_{\ell\le M} y(\ell)^2 \,, \tag{7.43}$$

where E_a denotes expectation in the r.v.s $a(\ell)$ only. By (7.36) we have

$$\left|\sum_{\ell\le M} y(\ell)\right| = |u_{N-1,M}(M)| \le 1$$

and $y(M)^2 = u_{N-1,M}(M)^4 \le 1$. Thus (7.43) implies

$$\mathsf{E}_a L_{N,M}(N)^2 \le \frac{K(\alpha)}{N} + \frac{L\beta^3}{\alpha^6 N}\sum_{\ell\le M-1} y(\ell)^2 \,. \tag{7.44}$$

To prove (7.42) we simply take expectation in (7.44), using that $M \le 3N$ and observing that $\mathsf{E} y(\ell)^2 = \mathsf{E} R_{N-1,M}(M-1)^2$ for $\ell \le M-1$. □

Proof of Theorem 7.4.1. To avoid trivial complications, we assume $\alpha \le 1$. Let us fix N, let us assume $M = \lfloor N(1+\alpha)\rfloor$, and, for $k \le N$ let us define

$$V(k) = \mathsf{E}\, R_{N-k,M-k}(M-k-1)^2 \,.$$

In this definition we assume that the values of $Z_{N-k,M'}$ that are relevant for the computation of $R_{N-k,M-k}$ have been computed with the parameter β replaced by the value β' such that $\beta'(N-k) = \beta N$. We observe that $M-k = \lfloor N(1+\alpha)-k\rfloor \ge \lfloor (N-k)(1+\alpha)\rfloor$ and $M-k \le 3(N-k)$. Combining Corollaries 7.4.6 and 7.4.4, implies that if $\beta'(N-k) = \beta N \ge 1$ and $\beta' \le \alpha/80$ we have

$$V(k) \le \frac{L\beta}{\alpha^6} V(k+1) + \frac{K(\alpha)}{N} \,. \tag{7.45}$$

Let us assume that $k \le N/2$, so that $b' \le 2b$. Then (7.45) holds whenever $\beta \le \alpha/160$. Thus if $L\beta/\alpha^6 \le 1/2$, $k \le N/2$ and $\beta N \ge 1$, we obtain

$$V(k) \le \frac{1}{2} V(k+1) + \frac{K(\alpha)}{N} \,.$$

Combining these relations yields

$$V(0) \le 2^{-k} V(k) + \frac{K(\alpha)}{N} \le 2^{-k+2} + \frac{K(\alpha)}{N}$$

since $V(k) \le 4$. Taking $k \simeq \log N$ proves (7.38), and (7.37) follows by (7.42). □

Theorem 7.4.7. *Under the conditions of Theorem 7.4.1, for $j \le M-1$, $i \le N-1$ we have*

$$\mathsf{E}\left(u_{N,M}(j) - u_{N,M-1}(j)\right)^2 \le \frac{K(\alpha)}{N} \tag{7.46}$$

$$\mathsf{E}\left(u_{N,M}(j) - u_{N-1,M}(j)\right)^2 \le \frac{K(\alpha)}{N} \tag{7.47}$$

$$\mathsf{E}\left(w_{N,M}(i) - w_{N,M-1}(i)\right)^2 \le \frac{K(\alpha)}{N} \tag{7.48}$$

$$\mathsf{E}\left(w_{N,M}(i) - w_{N-1,M}(i)\right)^2 \le \frac{K(\alpha)}{N} . \tag{7.49}$$

Proof. The proofs are similar, so we prove only (7.46). We can assume $j = M-1$. Using (7.29) and (7.35) we get

$$\begin{aligned} u_{N,M}(M-1) &= \frac{Z_{N,M}(\emptyset; M-1)}{Z_{N,M}} \\ &= \frac{Z_{N,M-2}}{Z_{N,M-1}} \left(\frac{1 + \sum_{k \le N} a(k,M)\, w_{N,M-2}(k)}{1 + \sum_{k \le N} a(k,M)\, w_{N,M-1}(k)} \right) . \end{aligned}$$

We observe the identity

$$L_{N,M}(i) = \frac{Z_{N,M-1}}{Z_{N,M}} \left(w_{N,M-1}(i) - w_{N,M}(i) \right) ,$$

which is obvious from (7.30). Using this identity for $M-1$ rather than M, we obtain

$$\begin{aligned} & u_{N,M}(M-1) - u_{N,M-1}(M-1) \\ &= \frac{Z_{N,M-2}}{Z_{N,M-1}} \left(\frac{1 + \sum_{k \le N} a(k,N)\, w_{N,M-2}(k)}{1 + \sum_{k \le N} a(k,N)\, w_{N,M-1}(k)} - 1 \right) \\ &= \frac{\sum_{k \le N} a(k,N)\, L_{N,M-1}(k)}{1 + \sum_{k \le N} a(k,N)\, w_{N,M-1}(k)} . \end{aligned}$$

Thus (7.47) follows from (7.37) and Lemma 7.4.3. □

We turn to the proof of Proposition 7.4.5, which occupies the rest of this section. It relies on the following probabilistic estimate.

Lemma 7.4.8. *Consider numbers* $0 \le u(\ell) \le 1$, *and let* $b = N^{-1} \sum_{\ell \le M} u(\ell)$. *Then if* $\beta N \ge 1$ *and* $\beta \le b/20$ *we have for* $k \le 8$ *that*

$$\mathsf{E}\left(\sum_{\ell \le M} a(\ell) u(\ell) \right)^{-k} \le \frac{L\beta^k}{b^k} , \tag{7.50}$$

where $a(\ell)$ *is as in (7.20).*

There is of course nothing magic about the number 8, this result is true for any other number (with a different condition on β). As the proof is tedious, it is postponed to the end of this section.

Proof of Proposition 7.4.5. First we reduce to the case $u(\ell) = u'(\ell)$ by using that $2cc' \le c^2 + c'^2$ for

$$c = \Big(\sum_{\ell \le M} a(\ell) u(\ell)\Big)^{-2} \quad ; \quad c' = \Big(\sum_{\ell \le M} a(\ell) u'(\ell)\Big)^{-2} .$$

Next, let $\dot{a}(\ell) = a(\ell) - \mathsf{E} a(\ell) = a(\ell) - \mathsf{E} a(1)$, so that

$$\sum_{\ell \le M} a(\ell) y(\ell) = \mathsf{E} a(1) \Big(\sum_{\ell \le M} y(\ell)\Big) + \sum_{\ell \le M} \dot{a}(\ell) y(\ell)$$

and since $\mathsf{E} a(1) \le 1/(\beta N)$,

$$\Big(\sum_{\ell \le M} a(\ell) y(\ell)\Big)^2 \le \frac{2}{\beta^2} \Big(\frac{1}{N} \sum_{\ell \le M} y(\ell)\Big)^2 + 2 \Big(\sum_{\ell \le M} \dot{a}(\ell) y(\ell)\Big)^2 .$$

Using (7.50) for $k = 4$, it suffices to prove that

$$\mathsf{E} \frac{\big(\sum_{\ell \le M} \dot{a}(\ell) y(\ell)\big)^2}{\big(\sum_{\ell \le M} a(\ell) u(\ell)\big)^4} \le \frac{L\beta^3}{b^6 N} \sum_{\ell \le M} y(\ell)^2 . \tag{7.51}$$

Expending the square in the numerator of the left-hand side, we see that it equals $\mathrm{I} + \mathrm{II}$, where

$$\mathrm{I} = \sum_{\ell' \le M} y(\ell')^2 \mathsf{E} \frac{\dot{a}(\ell')^2}{\big(\sum_{\ell \le M} a(\ell) u(\ell)\big)^4} \tag{7.52}$$

$$\mathrm{II} = \sum_{\ell_1 \ne \ell_2} y(\ell_1) y(\ell_2) \mathsf{E} \frac{\dot{a}(\ell_1) \dot{a}(\ell_2)}{\big(\sum_{\ell \le M} a(\ell) u(\ell)\big)^4} .$$

To bound the terms of I, let us set $S_{\ell'} = \sum_{\ell \ne \ell'} a(\ell) u(\ell)$, so

$$\mathsf{E} \frac{\dot{a}(\ell')^2}{\big(\sum_{\ell \le M} a(\ell) u(\ell)\big)^4} \le \mathsf{E} \frac{\dot{a}(\ell')^2}{S_{\ell'}^4} = \mathsf{E} \dot{a}(\ell')^2 \mathsf{E} \frac{1}{S_{\ell'}^4}$$

by independence. Now since $\sum_{\ell \le M} u(\ell) \ge 4$ and $u(\ell') \le 1$, we have

$$\sum_{\ell \ne \ell'} u(\ell) \ge \frac{3}{4} \sum_{\ell \le M} u(\ell) \ge \frac{3}{4} b , \tag{7.53}$$

so using (7.50) for $M - 1$ rather than M and $3b/4$ rather than b we get $\mathsf{E} S_{\ell'}^{-4} \le L\beta^4/b^4$; since $\mathsf{E} \dot{a}(\ell')^2 \le \mathsf{E} a(\ell')^2 \le 1/\beta N$, we have proved that, using that $b \le 1$ in the second inequality

$$\mathrm{I} \le \frac{L\beta^3}{Nb^4} \sum_{\ell \le M} y(\ell)^2 \le \frac{L\beta^3}{Nb^6} \sum_{\ell \le M} y(\ell)^2 \; .$$

To control the term II, let us set

$$S(\ell_1, \ell_2) = \sum_{\ell \ne \ell_1, \ell_2} a(\ell) u(\ell)$$

and

$$U = a(\ell_1)u(\ell_1) + a(\ell_2)u(\ell_2) \ge 0 \; .$$

Thus $\sum_{\ell \le M} a(\ell)u(\ell) = S(\ell_1, \ell_2) + U$. Since $U \ge 0$, a Taylor expansion yields

$$\frac{1}{\left(\sum_{\ell \le M} a(\ell)u(\ell)\right)^4} = \frac{1}{(S(\ell_1,\ell_2))^4} - \frac{4U}{S(\ell_1,\ell_2)^5} + \frac{\mathcal{R}}{S(\ell_1,\ell_2)^6} \tag{7.54}$$

where $|\mathcal{R}| \le 15U^2$. Since $S(\ell_1, \ell_2)$ is independent of $a(\ell_1)$ and $a(\ell_2)$, and since $\mathsf{E}\dot{a}(\ell_1)\dot{a}(\ell_2)U = 0$, multiplying (7.54) by $\dot{a}(\ell_1)\dot{a}(\ell_2)$ and taking expectation we get

$$\begin{aligned} \left| \mathsf{E} \frac{\dot{a}(\ell_1)\dot{a}(\ell_2)}{\left(\sum_{\ell \le M} a(\ell)u(\ell)\right)^4} \right| &\le \mathsf{E} \frac{15|\dot{a}(\ell_1)\dot{a}(\ell_2)|U^2}{S(\ell_1,\ell_2)^6} \\ &= 15\mathsf{E}(|\dot{a}(\ell_1)\dot{a}(\ell_2)|U^2)\mathsf{E}\frac{1}{S(\ell_1,\ell_2)^6} \; . \end{aligned}$$

Since $U^2 \le 2(a(\ell_1)^2 + a(\ell_2)^2)$ and $|\dot{a}(\ell_2)| \le 1$, independence implies

$$\mathsf{E}(|\dot{a}(\ell_1)\dot{a}(\ell_2)|U^2) \le 4\mathsf{E}(|\dot{a}(\ell_1)||\dot{a}(\ell_2)|a(\ell_2)^2) \le 4\mathsf{E}(|\dot{a}(\ell_1)|)\mathsf{E}a(\ell_2)^2 \; .$$

Now, $\mathsf{E}a(\ell)^2 \le 1/(2\beta N)$ and $\mathsf{E}|\dot{a}(\ell)| \le 2\mathsf{E}a(\ell) \le 2/(\beta N)$. Therefore we have

$$\mathsf{E}(|\dot{a}(\ell_1)\dot{a}(\ell_2)|U^2) \le \frac{L}{(\beta N)^2} \; .$$

We also have that $\mathsf{E}S(\ell_1,\ell_2)^{-6} \le L\beta^6/b^6$ by (7.50) (used for $k = 6$ and $M-2$ rather than M, and proceeding as in (7.53)). Thus

$$\begin{aligned} \mathrm{II} \le \frac{L\beta^4}{b^6N^2} \sum_{\ell_1 \ne \ell_2} |y(\ell_1)y(\ell_2)| &\le \frac{L\beta^4}{b^6N^2} \left(\sum_{\ell \le M} |y(\ell)| \right)^2 \\ &\le \frac{L\beta^4}{b^6N} \sum_{\ell \le M} y(\ell)^2 \; , \end{aligned}$$

and we conclude using that $\beta \le 1$. □

The following prepares the proof of Lemma 7.4.8.

Lemma 7.4.9. *If $\beta N \geq 1$ and $\lambda \geq 1$ we have*

$$\mathsf{E}\exp(-\lambda a(1)) \leq \exp\left(-\frac{\log \lambda}{2\beta N}\right) .$$

Proof. Assume first $\lambda \leq \exp \beta N$, so that $\log \lambda \leq \beta N$ and

$$\mathsf{P}(\lambda a(1) \geq 1) = \mathsf{P}(\exp \beta N X_1 \leq \lambda) = \mathsf{P}\left(X_1 \leq \frac{\log \lambda}{\beta N}\right) = \frac{\log \lambda}{\beta N} .$$

Thus, since $\exp(-x) \leq 1/2$ for $x \geq 1$, we have

$$\begin{aligned}\mathsf{E}\exp(-\lambda a(1)) &\leq 1 - \frac{1}{2}\mathsf{P}(\lambda a(1) \geq 1) \\ &\leq \exp\Big(-\frac{1}{2}\mathsf{P}(\exp(\beta N X_1) \leq \lambda)\Big) \\ &= \exp\Big(-\frac{\log \lambda}{2\beta N}\Big) .\end{aligned}$$

Consider next the case $\lambda \geq \exp \beta N$. Observe first that the function $\theta(x) = x/\log x$ increases for $x \geq e$ so that $\theta(\lambda) \geq \theta(\exp \beta N)$, i.e. $\lambda/\log(\lambda) \geq (\exp \beta N)/\beta N$, that is $\lambda \exp(-\beta N) \geq \log \lambda/\beta N$. Now, since $a(1) \geq \exp(-\beta N)$ we have

$$\mathsf{E}\exp(-\lambda a(1)) \leq \mathsf{E}\exp(-\lambda \exp(-\beta N)) \leq \exp\Big(-\frac{\log \lambda}{\beta N}\Big) . \qquad \square$$

Proof of Lemma 7.4.8. We use the inequality (A.8):

$$\mathsf{P}(Y \leq t) \leq (\exp \lambda t)\mathsf{E}\exp(-\lambda Y) \tag{7.55}$$

for $Y = \sum_{\ell \leq M} a(\ell)u(\ell)$ and any $\lambda \geq 0$. We have

$$\mathsf{E}\exp(-\lambda Y) = \mathsf{E}\exp\Big(-\lambda \sum_{\ell \leq M} a(\ell)u(\ell)\Big) = \prod_{\ell \leq M} \mathsf{E}\exp(-\lambda u_\ell a(\ell)) .$$

Since $u(\ell) \leq 1$, Hölder's inequality implies

$$\mathsf{E}\exp(-\lambda u_\ell a(\ell)) \leq \big(\mathsf{E}\exp(-\lambda a(\ell))\big)^{u(\ell)} = \big(\mathsf{E}\exp(-\lambda a(1))\big)^{u(\ell)} .$$

Therefore, assuming $\lambda \geq 1$, and using Lemma 7.4.9 in the second line,

$$\begin{aligned}\mathsf{E}\exp(-\lambda Y) &\leq \big(\mathsf{E}\exp(-\lambda a(1))\big)^{\sum_{\ell \leq M} u(\ell)} \\ &\leq \exp\Big(-\Big(\sum_{\ell \leq M} u(\ell)\Big)\frac{\log \lambda}{2\beta N}\Big) \\ &= \exp\left(-\frac{b \log \lambda}{2\beta}\right) ,\end{aligned} \tag{7.56}$$

using that $bN = \sum_{\ell \le M} u(\ell)$. Thus from (7.55) we get

$$\mathsf{P}\left(Y \le \frac{tb}{2e\beta}\right) \le \exp\left(-\frac{b}{2\beta}\left(\log\lambda - \frac{\lambda t}{e}\right)\right) . \tag{7.57}$$

For $t \le 1$, taking $\lambda = e/t$, and since then $\log\lambda - \lambda t/e = \log e/t - 1 = -\log t$, we get

$$\mathsf{P}\left(Y \le \frac{tb}{2e\beta}\right) \le t^{b/2\beta} .$$

Therefore whenever $t \ge 1$, the r.v. $X = 1/Y$ satisfies

$$\mathsf{P}\left(X \ge \frac{2te\beta}{b}\right) \le t^{-b/2\beta} . \tag{7.58}$$

Now we use (A.33) with $F(x) = x^k$ to get, making a change of variable in the second line,

$$\begin{aligned}\mathsf{E}X^k &= \int_0^\infty kt^{k-1}\mathsf{P}(X \ge t)\mathrm{d}t \\ &= \left(\frac{2e\beta}{b}\right)^k \int_0^\infty kt^{k-1}\mathsf{P}\left(X \ge \frac{2e\beta t}{b}\right)\mathrm{d}t .\end{aligned}$$

We bound $\mathsf{P}(X \ge 2e\beta t/b)$ by 1 for $t \le 1$ and using (7.58) for $t \ge 1$ to get

$$\mathsf{E}X^k \le \left(\frac{2e\beta}{b}\right)^k \left(1 + k\int_1^\infty t^{-b/(2\beta)+k-1}\mathrm{d}t\right) = \left(\frac{2e\beta}{b}\right)^k \left(1 + \frac{k}{b/(2\beta) - k}\right) ,$$

from which (7.50) follows since $k \le 8$ and $b/(2\beta) \ge 10$. □

Exercise 7.4.10. Prove that for a r.v. $Y \ge 0$ one has the formula

$$\mathsf{E}Y^{-k} = \frac{1}{(k-1)!}\int_0^\infty t^{k-1}\mathsf{E}\exp(-tY)\mathrm{d}t ,$$

and use it to obtain the previous bound on $\mathsf{E}X^k = \mathsf{E}Y^{-k}$ directly from (7.56).

7.5 Empirical Measures

Throughout the rest of this section, we assume the conditions of Theorem 7.4.1, that is, $\beta N \ge 1$, $M = \lfloor N(1+\alpha)\rfloor$ and $\beta \le \beta(\alpha)$.

Let us pursue our intuition that the sequence $(u_{N,M}(j))_{j \le M}$ looks like it is i.i.d. drawn out of a certain distribution. How do we find this distribution? The obvious candidate is the empirical measure

$$\mu_N = \frac{1}{M} \sum_{j \le M} \delta_{u_{N,M}(j)} \ . \tag{7.59}$$

We will also consider

$$\nu_N = \frac{1}{N} \sum_{i \le N} \delta_{w_{N,M}(i)} \ . \tag{7.60}$$

We recall the sequence $a(k) = \exp(-\beta N X_k)$, where (X_k) are i.i.d., uniform over $[0,1]$ and independent of the other sources of randomness. Consider the random measure $\overline{\mu}_N$ on $[0,1]$ given by

$$\overline{\mu}_N = \mathcal{L}_a\left(\frac{1}{1+\sum_{k\le N} a(k) w_{N,M}(k)}\right) ,$$

where $\mathcal{L}_a$ denotes the law in the randomness of the variables $a(k)$ with all the other sources of randomness fixed. Thus, for a continuous function f on $[0,1]$ we have

$$\int f \mathrm{d}\overline{\mu}_N = \mathsf{E}_a f\left(\frac{1}{1+\sum_{k\le N} a(k) w_{N,M}(k)}\right) ,$$

where E_a denotes expectation in the r.v.s $a(k)$ only. Consider the (non-random) measure $\mu'_N = \mathsf{E}\overline{\mu}_N$, so that

$$\int f \mathrm{d}\mu'_N = \mathsf{E} f\left(\frac{1}{1+\sum_{k\le N} a(k) w_{N,M}(k)}\right) .$$

In this section we shall show that $\mu_N \simeq \mu'_N$, and that, similarly, $\nu_N \simeq \nu'_N$ where

$$\int f \mathrm{d}\nu'_N = \mathsf{E} f\left(\frac{1}{\sum_{\ell\le M} a(\ell) u_{N,M}(\ell)}\right) .$$

In the next section we shall make precise the intuition that "ν'_N determines μ'_N" and "μ'_N determines ν'_N" to conclude the proof of Theorem 7.1.2.

It is helpful to consider an appropriate distance for probability measures. Given two probability measures μ, ν on $\mathbb{R}$, we consider the quantity

$$\Delta(\mu,\nu) = \inf \mathsf{E}(X-Y)^2 ,$$

where the infimum is over the pairs (X,Y) of r.v.s such that X has law μ and Y has law ν. The quantity $\Delta^{1/2}(\mu,\nu)$ is a distance. This statement is not obvious, but is proved in Section A.11, where the reader may find more information. This distance is called Wasserstein's distance between μ and ν. It is of course related to the transportation-cost distance considered in Chapter 6, but is more convenient here. Let us observe that since $\mathsf{E}(X-Y)^2 \ge (\mathsf{E}X - \mathsf{E}Y)^2$ we have

$$\left(\int x \mathrm{d}\mu(x) - \int x \mathrm{d}\nu(x)\right)^2 \le \Delta(\mu,\nu) \ . \tag{7.61}$$

Theorem 7.5.1. *The conditions of Theorem 7.4.1 imply*

$$\lim_{N\to\infty} \mathsf{E}\Delta(\mu_N, \mu'_N) = 0 \; ; \; \lim_{N\to\infty} \mathsf{E}\Delta(\nu_N, \nu'_N) = 0 \; . \tag{7.62}$$

We first collect some simple facts about Δ.

Lemma 7.5.2. *We have*

$$\Delta\left(\frac{1}{N}\sum_{i\le N}\delta_{x_i}\, ,\ \frac{1}{N}\sum_{i\le N}\delta_{y_i}\right) = \inf_\sigma \frac{1}{N}\sum_{i\le N}(x_i - y_{\sigma(i)})^2 \; , \tag{7.63}$$

where the infimum is over all permutations σ of $\{1,\dots,N\}$.

We will use this lemma when $x_i = w_{N,M}(i)$, and almost surely any two of these points are distinct. For this reason, we will give the proof only in the (easier) case where any two of the points x_i (resp. y_i) are distinct.

Proof. The inequality $\le$ should be obvious. To prove the converse inequality, we observe that if X has law $N^{-1}\sum_{i\le N}\delta_{x_i}$ and Y has law $N^{-1}\sum_{i\le N}\delta_{y_i}$, then

$$\mathsf{E}\,(X-Y)^2 = \sum_{i,j\le N} \mathsf{P}\,(X = x_i, Y = y_j)(x_i - y_j)^2 \; .$$

We observe that the bistochastic matrices are exactly the matrices $a_{ij} = N\mathsf{P}(X = x_i, Y = y_j)$. Thus the left-hand side of (7.63) is

$$\frac{1}{N}\inf \sum_{i,j\le N} a_{ij}(x_i - y_j)^2 \; ,$$

where the infimum is over all bistochastic matrices (a_{ij}). The infimum is attained at an extreme point, and it is a classical result ("Birkhoff's theorem") that this extreme point is a permutation matrix. □

Lemma 7.5.3. *Given numbers $w(k), w'(k) \ge 0$ we have*

$$\mathsf{E}\left(\frac{1}{1+\sum_{k\le N} a(k)w(k)} - \frac{1}{1+\sum_{k\le N} a(k)w'(k)}\right)^2 \le \frac{2}{\beta^2 N}\sum_{k\le N}(w(k) - w'(k))^2 \; . \tag{7.64}$$

Consequently

$$\Delta\left(\mathcal{L}\left(\frac{1}{1+\sum_{k\le N} a(k)w(k)}\right), \mathcal{L}\left(\frac{1}{1+\sum_{k\le N} a(k)w'(k)}\right)\right) \le \frac{2}{\beta^2 N}\sum_{k\le N}(w(k) - w'(k))^2 \; . \tag{7.65}$$

Proof. We use Lemma 7.4.3 together with the inequality

$$\left(\frac{1}{1+\sum_{k\le N} a(k)w(k)} - \frac{1}{1+\sum_{k\le M} a(k)w'(k)}\right)^2 \le \left(\sum_{k\le N} a(k)(w(k)-w'(k))\right)^2 .$$

□

The following fact is crucial.

Lemma 7.5.4. *For any continuous function f we have*

$$\lim_{N\to\infty} \mathsf{E}\left(f(u_{N,M}(M)) - \int f \mathrm{d}\overline{\mu}_N\right)\left(f(u_{N,M}(M-1)) - \int f \mathrm{d}\overline{\mu}_N\right) = 0 . \tag{7.66}$$

Proof. Recalling the numbers $a(k,\ell)$ of (7.6), let us consider

$$u = \frac{1}{1+\sum_{k\le N} a(k,M)w_{N,M-2}(k)} .$$

Using (7.27), (7.64) and (7.48) (with $M-1$ instead of M) we obtain

$$\mathsf{E}(u_{N,M}(M)-u)^2 \le \frac{K}{N} .$$

Exchanging the rôles of M and $M-1$ shows that if

$$u' = \frac{1}{1+\sum_{k\le N} a(k,M-1)w_{N,M-2}(k)}$$

we have

$$\mathsf{E}(u_{N,M}(M-1)-u')^2 = \mathsf{E}(u_{N,M}(M)-u)^2 \le \frac{K}{N} .$$

Therefore to prove (7.66) it suffices to prove that

$$\lim_{N\to\infty} \mathsf{E}\left(f(u) - \int f \mathrm{d}\overline{\mu}_N\right)\left(f(u') - \int f \mathrm{d}\overline{\mu}_N\right) = 0 . \tag{7.67}$$

Now by definition of $\overline{\mu}_N$ we have

$$\mathsf{E}\left(f(u) - \int f \mathrm{d}\overline{\mu}_N\right)\left(f(u') - \int f \mathrm{d}\overline{\mu}_N\right) = \mathsf{E}(f(u)-f(u_1))(f(u')-f(u_1')) ,$$

where

$$u_1 = \frac{1}{1+\sum_{k\le N} a(k)w_{N,M}(k)} ;\ u_1' = \frac{1}{1+\sum_{k\le N} a'(k)w_{N,M}(k)} ,$$

and where $a(k) = \exp(-\beta N X_k)$ and $a'(k) = \exp(-\beta N X'_k)$ are independent of all the other r.v.s involved. Let

$$u_2 = \frac{1}{1 + \sum_{k \le N} a(k) w_{N,M-2}(k)} \; ; \; u'_2 = \frac{1}{1 + \sum_{k \le M} a'(k) w_{N,M-2}(k)} \; .$$

Using again (7.64) and (7.48) we get

$$\mathsf{E}(u_1 - u_2)^2 \le \frac{K}{N} \; ; \; \mathsf{E}(u'_1 - u'_2)^2 \le \frac{K}{N} \; .$$

Therefore, to prove (7.67) it suffices to show that

$$\lim_{N \to \infty} \mathsf{E}(f(u) - f(u_2))(f(u') - f(u'_2)) = 0 \; .$$

Let us denote by E_a expectation only in the r.v.s $a(k), a'(k), a(k, M)$ and $a(k, M-1)$, which are probabilistically independent of the r.v.s $w_{N,M-2}(k)$. Then, by independence,

$$\mathsf{E}_a(f(u) - f(u_2))(f(u') - f(u'_2)) = (\mathsf{E}_a f(u) - \mathsf{E}_a f(u_2))(\mathsf{E}_a f(u') - \mathsf{E}_a f(u'_2)).$$

This is 0 because $\mathsf{E}_a f(u) = \mathsf{E}_a f(u_2)$, as is obvious from the definitions. □

Corollary 7.5.5. *For any continuous function f we have*

$$\lim_{N \to \infty} \mathsf{E}\left(\int f \mathrm{d}\mu_N - \int f \mathrm{d}\overline{\mu}_N\right)^2 = 0 \; . \tag{7.68}$$

Proof. We have

$$\int f \mathrm{d}\mu_N = \frac{1}{M} \sum_{\ell \le M} f(u_{N,M}(\ell))$$

so that, expanding the square and by symmetry

$$\mathsf{E}\left(\int f \mathrm{d}\mu_N - \int f \mathrm{d}\overline{\mu}_N\right)^2 = \frac{1}{M} \mathsf{E}\left(f(u_{N,M}(M)) - \int f \mathrm{d}\overline{\mu}_N\right)^2$$
$$+ \frac{M-1}{M} \mathsf{E}\left(f(u_{N,M}(M)) - \int f \mathrm{d}\overline{\mu}_N\right)\left(f(u_{N,M}(M-1)) - \int f \mathrm{d}\overline{\mu}_N\right) \; .$$

We conclude with Lemma 7.5.4. □

It is explained in Section A.11 why Wasserstein distance defines the weak topology on the set of probability measures on a compact space. Using (A.73) we see that (7.68) implies the following.

Corollary 7.5.6. *We have*

$$\lim_{N \to \infty} \mathsf{E}\Delta(\mu_N, \overline{\mu}_N) = 0 \; . \tag{7.69}$$

Lemma 7.5.7. *Consider an independent copy* $\widehat{\mu}_N$ *of the random measure* $\overline{\mu}_N$. *Then, recalling that* $\mu'_N = \mathsf{E}\overline{\mu}_N$, *we have*

$$\mathsf{E}\Delta(\overline{\mu}_N, \mu'_N) \le \mathsf{E}\Delta(\overline{\mu}_N, \widehat{\mu}_N)\ . \tag{7.70}$$

Proof. Let $\mathcal{C}$ be the class of pairs f, g of continuous functions such that

$$\forall x, y\ ,\ f(x) + g(y) \le (x-y)^2\ ,$$

so that by the duality formula (A.74) and since $\mu'_N = \mathsf{E}\overline{\mu}_N = \mathsf{E}\widehat{\mu}_N$,

$$\begin{aligned}\mathsf{E}\Delta(\overline{\mu}_N, \mu'_N) &= \mathsf{E} \sup_{(f,g)\in\mathcal{C}} \left(\int f \mathrm{d}\overline{\mu}_N + \mathsf{E}\int g \mathrm{d}\widehat{\mu}_N\right) \\ &\le \mathsf{E} \sup_{(f,g)\in\mathcal{C}} \left(\int f \mathrm{d}\overline{\mu}_N + \int g \mathrm{d}\widehat{\mu}_N\right) = \mathsf{E}\Delta(\overline{\mu}_N, \widehat{\mu}_N)\ ,\end{aligned}$$

using Jensen's inequality. □

Lemma 7.5.8. *Consider an independent copy* $\nu_N^{\sim}$ *of the random measure* ν_N *defined in (7.60). Then we have*

$$\mathsf{E}\Delta(\overline{\mu}_N, \widehat{\mu}_N) \le \frac{2}{\beta^2}\mathsf{E}\Delta(\nu_N, \nu_N^{\sim})\ .$$

Proof. Let $\nu_N^{\sim} = N^{-1}\sum_{k\le N}\delta_{w^{\sim}_{N,M}(k)}$, where $(w^{\sim}_{N,M}(k))_{k\le N}$ is an independent copy of the family $(w_{N,M}(k))_{k\le N}$. By Lemma 7.5.2 we can find a permutation σ with

$$\frac{1}{N}\sum_{k\le N}\big(w_{N,M}(k) - w^{\sim}_{N,M}(\sigma(k))\big)^2 = \Delta(\nu_N, \nu_N^{\sim})$$

and by Lemma 7.5.3 we get

$$\Delta(\overline{\mu}_N, \widehat{\mu}_N) \le \frac{2}{\beta^2}\Delta(\nu_N, \nu_N^{\sim}) \tag{7.71}$$

where

$$\widehat{\mu}_N = \mathcal{L}_a\left(\frac{1}{1+\sum_{k\le N} a(k) w^{\sim}_{N,M}(\sigma(k))}\right) = \mathcal{L}_a\left(\frac{1}{1+\sum_{k\le N} a(k) w^{\sim}_{N,M}(k)}\right)\ .$$

Taking expectation in (7.71) concludes the proof, since $\widehat{\mu}_N$ is an independent copy of $\widetilde{\mu}_N$. □

Let us observe the inequality

$$\Delta(\mu_1, \mu_2) \le 2(\Delta(\mu_1, \mu_3) + \Delta(\mu_3, \mu_2))\ , \tag{7.72}$$

which is a consequence of the fact that $\Delta^{1/2}$ is a distance.

Proposition 7.5.9. *We have*

$$\limsup_{N\to\infty} \mathsf{E}\Delta(\mu_N, \mu'_N) \le \frac{4}{\beta^2} \limsup_{N\to\infty} \mathsf{E}\Delta(\nu_N, \nu^\sim_N) \,. \tag{7.73}$$

Consequently, if $\mu^\sim_N$ denotes an independent copy of the random measure μ_N, we have

$$\limsup_{N\to\infty} \mathsf{E}\Delta(\mu_N, \mu^\sim_N) \le \frac{16}{\beta^2} \limsup_{N\to\infty} \mathsf{E}\Delta(\nu_N, \nu^\sim_N) \,. \tag{7.74}$$

Proof. Inequality (7.72) implies

$$\Delta(\mu_N, \mu'_N) \le 2\Delta(\mu_N, \overline{\mu}_N) + 2\Delta(\overline{\mu}_N, \mu'_N) \,.$$

Therefore (7.69) yields

$$\limsup_{N\to\infty} \mathsf{E}\Delta(\mu_N, \mu'_N) \le 2 \limsup_{N\to\infty} \mathsf{E}\Delta(\overline{\mu}_N, \mu'_N) \,.$$

By (7.70) and Lemma 7.5.8 this proves (7.73). To prove (7.74) we simply use (7.72) to write that

$$\Delta(\mu_N, \mu^\sim_N) \le 2\Delta(\mu_N, \mu'_N) + 2\Delta(\mu'_N, \mu^\sim_N) \,,$$

and we note that $\mathsf{E}\Delta(\mu'_N, \mu^\sim_N) = \mathsf{E}\Delta(\mu_N, \mu'_N)$. □

At this point we have done half of the work required to prove Theorem 7.5.1. The other half is as follows.

Proposition 7.5.10. *We have*

$$\limsup_{N\to\infty} \mathsf{E}\Delta(\nu_N, \nu'_N) \le \frac{L\beta^3}{\alpha^6} \limsup_{N\to\infty} \mathsf{E}\Delta(\mu_N, \mu^\sim_N) \tag{7.75}$$

and

$$\limsup_{N\to\infty} \mathsf{E}\Delta(\nu_N, \nu^\sim_N) \le \frac{L\beta^3}{\alpha^6} \limsup_{N\to\infty} \mathsf{E}\Delta(\mu_N, \mu^\sim_N) \,. \tag{7.76}$$

It is essential there to have a coefficient β^3 rather than β^2. Combining (7.76) and (7.74) shows that

$$\begin{aligned}\limsup_{N\to\infty} \mathsf{E}\Delta(\nu_N, \nu^\sim_N) &\le \frac{L\beta^3}{\alpha^6} \limsup_{N\to\infty} \mathsf{E}\Delta(\mu_N, \mu^\sim_N) \\ &\le \frac{L\beta^3}{\alpha^6} \frac{16}{\beta^2} \limsup_{N\to\infty} \mathsf{E}\Delta(\nu_N, \nu^\sim_N) \,,\end{aligned}$$

so that if $16L\beta/\alpha^6 < 1$ then

$$\limsup_{N\to\infty} \mathsf{E}\Delta(\nu_N, \nu^\sim_N) = \limsup_{N\to\infty} \mathsf{E}\Delta(\mu_N, \mu^\sim_N) = 0$$

and (7.73) and (7.75) prove Theorem 7.5.1.

The proof of Proposition 7.5.10 is similar to the proof of Proposition 7.5.9, using (7.41) rather than (7.40). Let us first explain the occurrence of the all important factor β^3 in (7.76).

Lemma 7.5.11. *Consider numbers $u(\ell), u'(\ell) \geq 0$ for $\ell \leq M$ and assume that $\sum_{\ell \leq M} u(\ell) = \sum_{\ell \leq M} u'(\ell) \geq N\alpha/2$. Then we have*

$$\mathsf{E}\left(\frac{1}{\sum_{\ell \leq M} a(\ell)u(\ell)} - \frac{1}{\sum_{\ell \leq M} a(\ell)u'(\ell)}\right)^2 \leq \frac{L\beta^3}{\alpha^6 N} \sum_{\ell \leq M} (u(\ell) - u'(\ell))^2 . \tag{7.77}$$

Consequently we have

$$\Delta\left(\mathcal{L}\left(\frac{1}{\sum_{\ell \leq M} a(\ell)u(\ell)}\right), \mathcal{L}\left(\frac{1}{\sum_{\ell \leq M} a(\ell)u'(\ell)}\right)\right) \leq \frac{L\beta^3}{\alpha^6 N} \sum_{\ell \leq M} (u(\ell) - u'(\ell))^2 . \tag{7.78}$$

Proof. We write

$$\left(\frac{1}{\sum_{\ell \leq M} a(\ell)u(\ell)} - \frac{1}{\sum_{\ell \leq M} a(\ell)u'(\ell)}\right)^2$$
$$\leq \frac{\left(\sum_{\ell \leq M} (u(\ell) - u'(\ell))a(\ell)\right)^2}{\left(\sum_{\ell \leq M} u(\ell)a(\ell)\right)^2 \left(\sum_{\ell \leq M} u'(\ell)a(\ell)\right)^2} ,$$

and we use (7.41) with $y(\ell) = u(\ell) - u'(\ell)$, so that $\sum_{\ell \leq M} y(\ell) = 0$. □

Consider the random measure $\overline{\nu}_N$ on $\mathbb{R}^+$ given by

$$\overline{\nu}_N = \mathcal{L}_a\left(\frac{1}{\sum_{\ell \leq N} a(\ell) u_{N,M}(\ell)}\right) ,$$

so that $\nu'_N = \mathsf{E}\overline{\nu}_N$. We denote by $\widehat{\nu}_N$ an independent copy of $\overline{\nu}_N$. We recall that $\mu_N^{\sim}$ denotes an independent copy of μ_N.

Lemma 7.5.12. *We have*

$$\mathsf{E}\Delta(\overline{\nu}_N, \widehat{\nu}_N) \leq \frac{L\beta^3}{\alpha^6} \mathsf{E}\Delta(\mu_N, \mu_N^{\sim}) .$$

Proof. Let $\mu_N^{\sim} = M^{-1} \sum_{\ell \leq M} \delta_{u^{\sim}_{N,M}(\ell)}$, where $(u^{\sim}_{N,M}(\ell))_{\ell \leq M}$ is an independent copy of the family $(u_{N,M}(\ell))_{\ell \leq M}$. By Lemma 7.5.2 we can find a permutation σ with

$$\frac{1}{M} \sum_{\ell \leq M} \left(u_{N,M}(\ell) - u^{\sim}_{N,M}(\sigma(\ell))\right)^2 = \Delta(\mu_N, \mu_N^{\sim}) .$$

The essential point now is that (7.26) yields

$$\sum_{\ell \leq M} u_{N,M}(\ell) = \sum_{\ell \leq M} u^{\sim}_{N,M}(\sigma(\ell)) = M - N \geq \alpha N/2 ,$$

so that we can use Lemma 7.5.11 to get

$$\Delta(\overline{\nu}_N, \widehat{\nu}_N) \le \frac{L\beta^3}{\alpha^6} \Delta(\mu_N, \mu_N^{\sim}) \tag{7.79}$$

where

$$\widehat{\nu}_N = \mathcal{L}_a\left(\frac{1}{\sum_{\ell \le M} a(\ell) u_{N,M}^{\sim}(\sigma(\ell))}\right) = \mathcal{L}_a\left(\frac{1}{\sum_{\ell \le M} a(\ell) u_{N,M}^{\sim}(\ell)}\right) .$$

Taking expectation in (7.79) concludes the proof, since $\widehat{\nu}_N$ is an independent copy of $\widetilde{\nu}_N$. □

The rest of the arguments in the proof of Proposition 7.5.10 is very similar to the arguments of Proposition 7.5.9. One extra difficulty is that the distributions ν_N (etc.) no longer have compact support. This is bypassed by a truncation argument. Indeed, it follows from (7.28) and (7.50) that

$$\mathsf{E} w_{N,M}^4(i) \le K(\alpha) .$$

If $b \ge 0$ is a truncation level, the quantities $w_{N,M,b}(i) := \min(w_{N,M}(i), b)$ satisfy

$$\mathsf{E}(w_{N,M}(i) - w_{N,M,b}(i))^2 \le \mathsf{E}\big(w_{N,M}^2(i) \mathbf{1}_{\{w_{N,M}(i) \ge b\}}\big) \le \frac{K(\alpha)}{b^2} .$$

If we define $\nu_{N,b} = N^{-1} \sum_{i \le N} \delta_{w_{N,M,b}(i)}$, then

$$\Delta(\nu_{N,b}, \nu_N) \le \frac{1}{N} \sum_{i \le N} (w_{N,M}(i) - w_{N,M,b}(i))^2$$

so that

$$\mathsf{E}\Delta(\nu_{N,b}, \nu_N) \le \frac{K(\alpha)}{b^2} , \tag{7.80}$$

and using such a uniformity, rather than (7.75) it suffices to prove for each b the corresponding result when in the left-hand side "everything is truncated at level b". More specifically, defining $\nu'_{N,b}$ by

$$\int f \mathrm{d}\nu'_{N,b} = \mathsf{E} f\left(\min\left(b, \frac{1}{\sum_{\ell \le M} a(\ell) u_{N,M}(\ell)}\right)\right) ,$$

one proves that

$$\limsup_{N \to \infty} \mathsf{E}\Delta(\nu_{N,b}, \nu'_{N,b}) \le \frac{L\beta^3}{\alpha^6} \limsup_{N \to \infty} \mathsf{E}\Delta(\mu_N, \mu_N^{\sim}) ,$$

and one uses that (7.80) implies

$$\limsup_{N \to \infty} \mathsf{E}\Delta(\nu_N, \nu'_N) \le \limsup_{N \to \infty} \mathsf{E}\Delta(\nu_{N,b}, \nu'_{N,b}) + \frac{K(\alpha)}{b^2} .$$

The details are straightforward.

7.6 Operators

The definition of the operators A and B given in (7.9) and (7.10) is pretty, but it does not reflect the property we need. The fundamental property of the operator A is that if the measure $M^{-1}\sum_{\ell\le M}\delta_{u(\ell)}$ approaches the measure μ, the law of $\left(\sum_{\ell\le M} a_N(\ell)u(\ell)\right)^{-1}$ approaches $A(\mu)$, where $a_N(\ell) = \exp(-N\beta X_\ell)$, $M/N \simeq 1+\alpha$, and where of course the r.v.s $(X_\ell)_{\ell\ge 1}$ are i.i.d. uniform over $[0,1]$. Since the description of A given in (7.9) will not be needed, its (non-trivial) equivalence with the definition we will give below in Proposition 7.6.2 will be left to the reader.

In order to prove the existence of the operator A, we must prove that if two measures

$$\frac{1}{M}\sum_{\ell\le M}\delta_{u(\ell)} \quad \text{and} \quad \frac{1}{M'}\sum_{\ell\le M'}\delta_{u'(\ell)}$$

both approach μ, and if $M/N \simeq M'/N'$, then

$$\mathcal{L}\left(\frac{1}{\sum_{\ell\le M} a_N(\ell)u(\ell)}\right) \simeq \mathcal{L}\left(\frac{1}{\sum_{\ell\le M'} a_{N'}(\ell)u'(\ell)}\right) .$$

This technical fact is contained in the following estimate.

Proposition 7.6.1. *Consider a number $\alpha > 0$. Consider integers M, N, M', N' with $N \le M \le 2N$, $N' \le M' \le 2N'$ and numbers $0 \le u(\ell) \le 1$ for $\ell \le M$, numbers $0 \le u'(\ell) \le 1$ for $\ell \le M'$. Let*

$$\eta = \frac{1}{M}\sum_{\ell\le M}\delta_{u(\ell)} \ ; \ \eta' = \frac{1}{M'}\sum_{\ell\le M'}\delta_{u'(\ell)} .$$

Assume that $\int x\mathrm{d}\eta(x) \ge \alpha/4$ and $\int x\mathrm{d}\eta'(x) \ge \alpha/4$. Assume that $\beta N \ge 1, \beta N' \ge 1$ and $\beta \le \alpha/80$. Then, with $a_N(\ell) = \exp(-\beta N X_\ell)$ as above, we have

$$\begin{aligned}\Delta\left(\mathcal{L}\left(\frac{1}{\sum_{\ell\le M} a_N(\ell)u(\ell)}\right), \mathcal{L}\left(\frac{1}{\sum_{\ell\le M'} a_{N'}(\ell)u'(\ell)}\right)\right) \\ \le K(\alpha)\left(\frac{1}{N}+\frac{1}{N'}+\left|\frac{M}{N}-\frac{M'}{N'}\right|\right) \\ + \frac{L\beta^3}{\alpha^6}\Delta(\eta,\eta') + \frac{L\beta^2}{\alpha^4}\left(\int x\mathrm{d}\eta(x) - \int x\mathrm{d}\eta'(x)\right)^2 . \qquad (7.81)\end{aligned}$$

Let us state an important consequence.

Proposition 7.6.2. *Given a number $\alpha > 0$ there exists a number $\beta(\alpha) > 0$ with the following property. If $\beta \le \beta(\alpha)$ and if μ is a probability measure on*

$[0,1]$ *with* $\int x \mathrm{d}\mu(x) \geq \alpha/4$, *there exists a unique probability measure* $A(\mu)$ *on* $\mathbb{R}^+$ *with the following property. Consider numbers* $0 \leq u(\ell) \leq 1$ *for* $\ell \leq M$, *and set*

$$\eta = \frac{1}{M} \sum_{\ell \leq M} \delta_{u(\ell)} .$$

Then

$$\begin{aligned} \Delta\left(A(\mu), \mathcal{L}\Big(\frac{1}{\sum_{\ell \leq M} a_N(\ell) u(\ell)}\Big)\right) &\leq K(\alpha)\left(\frac{1}{N} + \left|\frac{M}{N} - (1+\alpha)\right|\right) \\ &+ \frac{L\beta^2}{\alpha^4}\left(\int x \mathrm{d}\mu(x) - \int x \mathrm{d}\eta(x)\right)^2 \\ &+ \frac{L\beta^3}{\alpha^6} \Delta(\mu, \eta) . \end{aligned} \tag{7.82}$$

Moreover, if μ' *is another probability measure and if* $\int x \mathrm{d}\mu'(x) \geq \alpha/4$, *we have*

$$\Delta(A(\mu), A(\mu')) \leq \frac{L\beta^2}{\alpha^4}\left(\int x \mathrm{d}\mu(x) - \int x \mathrm{d}\mu'(x)\right)^2 + \frac{L\beta^3}{\alpha^6} \Delta(\mu, \mu') . \tag{7.83}$$

A little bit of measure-theoretic technique is required again here, because we are dealing with probability measures that are not supported by a compact interval. In the forthcoming lemma, there is really nothing specific about the power 4.

Lemma 7.6.3. *Given a number* C, *consider the set* $D(C)$ *of probability measures* θ *on* $\mathbb{R}^+$ *that satisfy* $\int_0^\infty x^4 \mathrm{d}\theta(x) \leq C$. *Then* $D(C)$ *is a compact metric space for the distance* Δ.

Proof. The proof uses a truncation argument similar to the one given at the end of the proof of Proposition 7.5.10. Given a number $b > 0$ and a probability measure θ in $D(C)$ we define the truncation θ^b as the image of θ under the map $x \mapsto \min(x, b)$. In words, all the mass that θ gives to the half-line $[b, \infty[$ is pushed to the point b. Then we have

$$\Delta(\theta, \theta^b) \leq \int_0^\infty (x - \min(x,b))^2 \mathrm{d}\theta(x) \leq \int_b^\infty x^2 \mathrm{d}\theta(x) \leq \frac{C}{b^2} . \tag{7.84}$$

Consider now a sequence $(\theta_n)_{n \geq 1}$ in $D(C)$. We want to prove that it has a subsequence that converges for the distance Δ. Since for each b the set of probability measures on the interval $[0,b]$ is compact for the distance Δ (as is explained in Section A.11), we assume, by taking a subsequence if necessary, that for each integer m the sequence $(\theta_n^m)_{n \geq 1}$ converges for Δ to a certain probability measure λ_m. Next we show that there exists a probability measure λ in $D(C)$ such that $\lambda_m = \lambda^m$ for each m. This is simply because if $m' < m$

then $\lambda_m^{m'} = \lambda_{m'}$ (the "pieces fit together") and because $\int_0^\infty x^4 \mathrm{d}\lambda_m(x) \le C$ for each m. Now, for each m we have $\lim_{n\to\infty} \Delta(\theta_n^m, \lambda^m) = 0$, and (7.84) and the triangle inequality imply that $\lim_{n\to\infty} \Delta(\theta_n, \lambda) = 0$. □

Proof of Proposition 7.6.2. The basic idea is to define $A(\mu)$ "as the limit" of the law λ of $\left(\sum_{\ell\le M} a_N(\ell)u(\ell)\right)^{-1}$ as $M^{-1}\sum_{\ell\le M}\delta_{u(\ell)} \to \mu$, $M, N \to \infty$, $M/N \to (1+\alpha)$. We note that by (7.50) used for $k = 8$, whenever $\sum_{\ell\le M} u(\ell) \ge \alpha N/8$, (and $\beta < \beta(\alpha)$) we have $\int x^4 \mathrm{d}\lambda(x) \le L$. Thus, recalling the notation of Lemma 7.6.3, we have $\lambda \in D(L)$, a compact set, and therefore the family of these measures has a cluster point $A(\mu)$, and (7.82) holds by continuity. Moreover (7.83) is a consequence of (7.82) and continuity (and shows that the cluster point $A(\mu)$ is in fact unique). □

We recall the probability measures $\mu_N, \nu_N, \nu'_N, \mu'_N$ of Section 7.5.

Proposition 7.6.4. *We have*

$$\lim_{N\to\infty} \Delta(\nu'_N, A(\mu'_N)) = 0 \,. \tag{7.85}$$

Proof. First we recall that by (7.26) we have

$$\int x \mathrm{d}\mu_N(x) = \frac{1}{M}\sum_{\ell\le M} u_{N,M}(\ell) = \frac{M-N}{M} \ge \frac{\alpha}{2}$$

for $M = \lfloor N(1+\alpha)\rfloor$ and N large. Since Theorem 7.5.1 asserts that $\mathsf{E}\Delta(\mu'_N, \mu_N) \to 0$, (7.61) implies that

$$\mathsf{E}\left(\int x\mathrm{d}\mu_N(x) - \int x\mathrm{d}\mu'_N(x)\right)^2 \to 0$$

and thus $\int x\mathrm{d}\mu'_N(x) \ge \alpha/4$ for N large. Therefore we can use (7.82) for $\mu = \mu'_N$ and $\eta = \mu_N$ to get (using (7.61) again)

$$\begin{aligned} &\Delta\left(A(\mu'_N), \mathcal{L}_a\left(\frac{1}{\sum_{\ell\le M} a_N(\ell)u_{N,M}(\ell)}\right)\right) \\ &\le \frac{K(\alpha)}{N} + L\left(\frac{\beta^2}{\alpha^4} + \frac{\beta^3}{\alpha^6}\right)\Delta(\mu'_N, \mu_N) \,. \end{aligned} \tag{7.86}$$

The expectation of the right-hand side goes to zero as $N \to \infty$ by Theorem 7.5.1. Since by definition

$$\nu'_N = \mathsf{E}\mathcal{L}_a\left(\frac{1}{\sum_{\ell\le M} a_N(\ell)u_{N,M}(\ell)}\right),$$

taking expectation in (7.86) and using Jensen's inequality as in (7.70) completes the proof. □

Proposition 7.6.4 is of course only half of the work because we also have to define the operators B. These operators B have the following defining property.

Proposition 7.6.5. *To each probability measure ν on $\mathbb{R}^+$ we can attach a probability measure $B(\nu)$ on $[0,1]$ with the following property. Consider numbers $w(k) \geq 0$ for $k \leq N$, and let*

$$\eta = \frac{1}{N} \sum_{k \leq N} \delta_{w(k)} \; .$$

Then

$$\Delta\left(B(\nu), \mathcal{L}\Big(\frac{1}{1+\sum_{k\leq N} a_N(k) w(k)}\Big)\right) \leq \frac{K}{N} + \frac{L}{\beta^2} \Delta(\nu, \eta) \; . \tag{7.87}$$

Moreover

$$\Delta(B(\nu), B(\nu')) \leq \frac{L}{\beta^2} \Delta(\nu, \nu') \; . \tag{7.88}$$

Proof. Similar, but simpler than the proof of Proposition 7.6.2. □

Proposition 7.6.6. *We have*

$$\lim_{N\to\infty} \Delta(\mu'_N, B(\nu'_N)) = 0 \; . \tag{7.89}$$

Proof. Similar (but simpler) than the proof of (7.85). □

Proof of Theorem 7.1.2. It follows from the definition of ν'_N and (7.50) that $\int x^4 \mathrm{d}\nu'_N(x) \leq L$, so that, recalling the set $D(L)$ of Lemma 7.6.3, we have $\nu'_N \in D(L)$. Since μ'_N lives on $[0,1]$, we can find a subsequence of the sequence (μ'_N, ν'_N) that converges for Δ to a pair (μ, ν). Using (7.85) and (7.89) we see that this pair satisfies the relations (7.11):

$$\int x \mathrm{d}\mu(x) = \frac{\alpha}{1+\alpha} \; ; \; \mu = B(\nu) \; , \; \nu = A(\mu) \; . \tag{7.90}$$

The equations (7.90) have a unique solution. Indeed, if (μ', ν') is another solution (7.83) implies

$$\Delta(\nu, \nu') \leq \frac{L\beta^3}{\alpha^6} \Delta(\mu', \mu)$$

and by (7.88) we have

$$\Delta(\mu, \mu') \leq \frac{L}{\beta^2} \Delta(\nu, \nu')$$

so that

$$\Delta(\mu, \mu') \leq \frac{L\beta}{\alpha^6} \Delta(\mu, \mu')$$

and $\Delta(\mu, \mu') = 0$ if $L\beta/\alpha^6 < 1$. Let us stress the miracle here. The condition (7.26) forces the relation $\int x \mathrm{d}\mu(x) = \alpha/(1+\alpha)$, and this neutralizes the first term on the right-hand side of (7.83). This term is otherwise devastating, because the coefficient $L\beta^2/\alpha^4$ does not compensate the coefficient L/β^2 of (7.88).

Since the pair (μ, ν) of (7.90) is unique, we have in fact that $\mu = \lim \mu'_N$, $\nu = \lim \nu'_N$. On the other hand, by definition of μ_N we have $\mathsf{E}\mu_N = \mathcal{L}(u_{N,M}(M))$, so Jensen's inequality implies as in (7.70) that

$$\Delta(\mathcal{L}(u_{N,M}(M)), \mu'_N) \le \mathsf{E}\Delta(\mu_N, \mu'_N) ,$$

so $\lim_{N\to\infty} \mathcal{L}(u_{N,M}(M)) = \mu$ by (7.62). Similarly $\lim_{N\to\infty} \mathcal{L}(w_{N,M}(N)) = \nu$. □

We turn to the proof of Proposition 7.6.1. Let us start by a simple observation.

Proposition 7.6.7. *The bound (7.81) holds when $M = M'$.*

Proof. Without loss of generality we assume that $N' \le N$. Let $S = \sum_{\ell\le M} a_N(\ell)u(\ell)$ and $S' = \sum_{\ell\le M} a_{N'}(\ell)u'(\ell)$. Then

$$\Delta\left(\mathcal{L}\left(\frac{1}{S}\right), \mathcal{L}\left(\frac{1}{S'}\right)\right) \le \mathsf{E}\left(\frac{1}{S} - \frac{1}{S'}\right)^2 = \mathsf{E}\frac{(S-S')^2}{S^2S'^2} \le \mathrm{I} + \mathrm{II} \tag{7.91}$$

where

$$\mathrm{I} = 2\mathsf{E}\,\frac{\left(\sum_{\ell\le M}(a_N(\ell) - a_{N'}(\ell))u'(\ell)\right)^2}{S^2S'^2} ;$$

$$\mathrm{II} = 2\mathsf{E}\,\frac{\left(\sum_{\ell\le M} a_N(\ell)(u(\ell) - u'(\ell))\right)^2}{S^2S'^2} .$$

We observe since $N' \le N$ that $a'_N(\ell) \ge a_N(\ell)$, so that

$$S' \ge S^\sim := \sum_{\ell\le M} a_N(\ell)u'(\ell) ,$$

and

$$\mathrm{II} \le 2\mathsf{E}\,\frac{\left(\sum_{\ell\le M} a_N(\ell)(u(\ell) - u'(\ell))\right)^2}{S^2S^{\sim 2}} .$$

To bound this quantity we will use the estimate (7.41). The relations $\int x\mathrm{d}\eta(x) \ge \alpha/4$ and $\int x\mathrm{d}\eta'(x) \ge \alpha/4$ mean that $\sum_{\ell\le M} u(\ell) \ge \alpha M/4 \ge \alpha N/4$ and $\sum_{\ell\le M} u'(\ell) \ge \alpha M/4 \ge \alpha N/4$. Thus in (7.41) we can take $b = \alpha/4$. This estimate then yields

$$\mathrm{II} \le \frac{L\beta^2}{\alpha^4}\left(\frac{M}{N}\right)^2\left(\int x\mathrm{d}\eta(x) - \int x\mathrm{d}\eta'(x)\right)^2 + \frac{L\beta^3}{\alpha^6}\frac{1}{N}\sum_{\ell\le M}(u(\ell) - u'(\ell))^2 . \tag{7.92}$$

We can assume from Lemma 7.5.2 that we have reordered the terms $u'(\ell)$ so that $M^{-1}\sum_{\ell\le M}(u(\ell)-u'(\ell))^2 \le \Delta(\eta,\eta')$, and then the bound (7.92) is as desired, since $M \le 2N$.

To control the term I, we first note that $0 \le a_{N'}(\ell) - a_N(\ell) \le 1$ since $N' \le N$; and $\sum_{\ell\le M}(a_{N'}(\ell)-a_N(\ell))u(\ell) \le M$ since $0 \le u'(\ell) \le 1$. Therefore

$$\mathrm{I} \le 2M \sum_{\ell\le M} \mathsf{E}\frac{a_{N'}(\ell)-a_N(\ell)}{S^2 S'^2} .$$

We control this term with the same method that we used to control the term (7.52). Namely, we define $S_\ell = \sum_{\ell'\ne\ell} a_N(\ell')u(\ell')$ and S'_ℓ similarly, and we write, using independence and the Cauchy-Schwarz inequality that

$$\begin{aligned}\mathsf{E}\,\frac{a_{N'}(\ell)-a_N(\ell)}{S^2S'^2} &\le \mathsf{E}\,\frac{a_{N'}(\ell)-a_N(\ell)}{S_\ell^2 S_\ell'^2}\\ &\le \mathsf{E}(a_{N'}(\ell)-a_N(\ell))\Big(\mathsf{E}\,\frac{1}{S_\ell^4}\Big)^{1/2}\Big(\mathsf{E}\,\frac{1}{S_\ell'^4}\Big)^{1/2} .\end{aligned}$$

Using (7.50), and since $\sum_{\ell\ne\ell'} u(\ell) \ge N\alpha/4 - 1 \ge N\alpha/8$ because $N\beta \ge 1$ and $\beta \le \alpha/80$, we get

$$\Big(\mathsf{E}\,\frac{1}{S_\ell^4}\Big)^{1/2} \le K(\alpha)\beta^2 ,$$

and similarly for S'_ℓ. Using (7.39) for $p=1$, we obtain

$$\mathsf{E}(a_{N'}(\ell)-a_N(\ell)) \le \frac{L}{\beta}\Big(\frac{1}{N'}-\frac{1}{N}\Big) .$$

The result follows. □

The main difficulty in the proof of Proposition 7.6.1 is to find how to relate the different values M and M'. Given a sequence $(u(\ell))_{\ell\le M}$ and an integer M', consider the sequence $(u^{\sim}(\ell))_{\ell\le MM'}$ that is obtained by repeating each term $u(\ell)$ exactly M' times.

Proposition 7.6.8. *We have*

$$\Delta\left(\mathcal{L}\Big(\frac{1}{\sum_{\ell\le M} a_N(\ell)u(\ell)}\Big), \mathcal{L}\Big(\frac{1}{\sum_{\ell\le MM'} a_{NM'}(\ell)u^{\sim}(\ell)}\Big)\right) \le \frac{K}{N} . \tag{7.93}$$

Proof of Proposition 7.6.1. The meaning of (7.93) is that within a small error (as in (7.81)) we can replace M by MM' and N by NM'. Similarly, we replace M' by MM' and N' by $N'M$, so we have reduced the proof to the case $M = M'$ of Proposition 7.6.7 (using that $\Delta^{1/2}$ is a distance). □

The proof of Proposition 7.6.8 relies on the following.

Lemma 7.6.9. *Consider independent r.v.s X_ℓ, X, uniform over $[0,1]$. Consider an integer $R \ge 1$, a number $\gamma \ge 2$ and the r.v.s*

$$a = \exp(-\gamma X)\,;\; a' = \sum_{\ell \le R} \exp(-\gamma R\, X_\ell)\,.$$

Then we can find a pair of r.v.s (Y, Y') such that Y has the same law as the r.v. a and Y' has the same law as the r.v. a' with

$$\mathsf{E}|Y - Y'| \le \frac{L}{\gamma^2}\,,\; \mathsf{E}(Y - Y')^2 \le \frac{L}{\gamma^2}\,. \tag{7.94}$$

Proof of Proposition 7.6.8. We use Lemma 7.6.9 for $\gamma = \beta N$, $R = M'$. Consider independent copies (Y_ℓ, Y'_ℓ) of the pair (Y, Y'). It should be obvious from the definition of the sequence $u^\sim(\ell)$ that $S' := \sum_{\ell \le M} Y'_\ell u(\ell)$ equals $\sum_{\ell \le MM'} a_{MM'}(\ell) u^\sim(\ell)$ in distribution. Writing $S = \sum_{\ell \le M} Y_\ell u(\ell)$, the left-hand side of (7.93) is

$$\begin{aligned}\Delta\left(\mathcal{L}\left(\frac{1}{S}\right), \mathcal{L}\left(\frac{1}{S'}\right)\right) \le \mathsf{E}\left(\frac{1}{S} - \frac{1}{S'}\right)^2 &= \mathsf{E}\,\frac{\left(\sum_{\ell \le M}(Y_\ell - Y'_\ell)u(\ell)\right)^2}{S^2 S'^2}\,,\\ &\le \mathsf{E}\,\frac{\left(\sum_{\ell \le M}|Y_\ell - Y'_\ell|\right)^2}{S^2 S'^2}\,.\end{aligned}$$

We expand the square, and we use (7.94) for $\gamma = \beta N$ and one more time the method used to control (7.52) to find that this is $\le K(\alpha)/N$. □

Proof of Lemma 7.6.9. Given any two r.v.s $a, a' \ge 0$, there is a canonical way to construct a coupling of them. Consider the function Y on $[0,1]$ given by

$$Y(x) = \inf\{t\,;\; \mathsf{P}(a \ge t) \le x\}\,.$$

The law of Y under Lebesgue's measure is the law of a. Indeed the definition of $Y(x)$ shows that

$$\begin{aligned}\mathsf{P}(a \ge y) > x &\Rightarrow Y(x) > y\\ \mathsf{P}(a \ge y) < x &\Rightarrow Y(x) < y\,,\end{aligned}$$

so that if λ denotes Lebesgue measure, we have $\lambda(\{Y(x) \ge y\}) = \mathsf{P}(a \ge y)$. Moreover "the graph of Y is basically obtained from the graph of the function $t \mapsto \mathsf{P}(a \ge t)$ by making a symmetry around the diagonal". Define Y' similarly. The pair (Y, Y') is the pair we look for, although it will require some work to prove this. First we note that

$$\mathsf{E}|Y - Y'| = \int_0^1 |Y(x) - Y'(x)|\mathrm{d}x\,.$$

This is the area between the graphs of Y of Y', and also the area between the graphs of the functions $t \mapsto \mathsf{P}(a \ge t)$ and $t \mapsto \mathsf{P}(a' \ge t)$ because these two areas are exchanged by symmetry around the diagonal (except maybe for their boundary). Therefore

$$\mathsf{E}|Y - Y'| = \int_0^\infty |\mathsf{P}(a \geq t) - \mathsf{P}(a' \geq t)| \mathrm{d}t \ .$$

The rest of the proof consists in elementary (and very tedious) estimates of this quantity when a and a' are as in Lemma 7.6.9. For $t \leq 1$ we have

$$\mathsf{P}(a \geq t) = \mathsf{P}(\exp(-\gamma X) \geq t) = \mathsf{P}\left(X \leq \frac{1}{\gamma} \log \frac{1}{t}\right) = \min\left(1, \frac{1}{\gamma} \log \frac{1}{t}\right) ,$$

and similarly

$$\mathsf{P}(\exp(-\gamma R X_\ell) \geq t) = \min\left(1, \frac{1}{\gamma R} \log \frac{1}{t}\right) .$$

Since $a' \geq t$ as soon as one of the summands $\exp(-\gamma R X_\ell)$ exceeds t, independence implies

$$\mathsf{P}(a' \geq t) \geq 1 - \left(1 - \min\left(1, \frac{1}{\gamma R} \log \frac{1}{t}\right)\right)^R := \psi(t) .$$

Since $(1-x)^R \geq 1 - Rx$ for $x \geq 0$, we have

$$\psi(t) \leq R \min\left(1, \frac{1}{\gamma R} \log \frac{1}{t}\right) = \min\left(R, \frac{1}{\gamma} \log \frac{1}{t}\right) ,$$

and since $\psi(t) \leq 1$, we have in fact

$$\psi(t) \leq \min\left(1, \frac{1}{\gamma} \log \frac{1}{t}\right) = \mathsf{P}(a \geq t) .$$

We note that

$$x \geq 0 \Rightarrow (1-x)^R \leq e^{-Rx} \leq 1 - Rx + \frac{R^2 x^2}{2} .$$

Using this for

$$x = \min\left(1, \frac{1}{R\gamma} \log \frac{1}{t}\right)$$

this yields that

$$\psi(t) = 1 - (1-x)^R \geq Rx - \frac{R^2 x^2}{2}$$

and

$$0 \leq \mathsf{P}(a \geq t) - \psi(t) \leq \min\left(1, \frac{1}{\gamma} \log \frac{1}{t}\right) - Rx + \frac{R^2 x^2}{2} .$$

Since

$$\min\left(1, \frac{1}{\gamma}\log\frac{1}{t}\right) \leq Rx \leq \frac{1}{\gamma}\log\frac{1}{t} ,$$

we have proved that

$$0 \leq \mathsf{P}(a \geq t) - \psi(t) \leq \frac{1}{2}\left(\frac{1}{\gamma}\log\frac{1}{t}\right)^2 . \tag{7.95}$$

For a real number y we write $y^+ = \max(y, 0)$, so that $|y| = -y + 2y^+$. We use this relation for $y = \mathsf{P}(a \geq t) - \mathsf{P}(a' \geq t)$, so that since $\mathsf{P}(a' \geq t) \geq \psi(t)$ we obtain

$$y^+ \leq (\mathsf{P}(a \geq t) - \psi(t))^+ = \mathsf{P}(a \geq t) - \psi(t) ,$$

and

$$|\mathsf{P}(a \geq t) - \mathsf{P}(a' \geq t)| \leq \mathsf{P}(a' \geq t) - \mathsf{P}(a \geq t) + 2(\mathsf{P}(a \geq t) - \psi(t)) . \tag{7.96}$$

Since $a \leq 1$, for $t > 1$ we then have

$$|\mathsf{P}(a \geq t) - \mathsf{P}(a' \geq t)| = \mathsf{P}(a' \geq t) = \mathsf{P}(a' \geq t) - \mathsf{P}(a \geq t) . \tag{7.97}$$

Using (7.96) for $t \leq 1$ and (7.97) for $t > 1$ we obtain, using (7.95) in the second inequality,

$$\begin{aligned}\int_0^\infty |\mathsf{P}(a \geq t) - \mathsf{P}(a' \geq t)|\,\mathrm{d}t &\leq 2\int_0^1 (\mathsf{P}(a \geq t) - \psi(t))\,\mathrm{d}t \\ &\quad + \int_0^\infty \mathsf{P}(a' \geq t)\,\mathrm{d}t - \int_0^\infty \mathsf{P}(a \geq t)\,\mathrm{d}t \\ &\leq \frac{L}{\gamma^2} + \mathsf{E}a' - \mathsf{E}a .\end{aligned}$$

Finally we use that by (7.39) we have $|\mathsf{E}a - \mathsf{E}a'| \leq L/\gamma^2$, and this concludes the proof that $\mathsf{E}\,|Y - Y'| \leq L/\gamma^2$.

We turn to the control of $\mathsf{E}(Y - Y')^2$. First, we observe that

$$\mathsf{E}(Y - Y')^2 \leq 2\mathsf{E}(Y - \min(Y', 2))^2 + 2\mathsf{E}(\min(Y', 2) - Y')^2 .$$

Now, since $Y \leq 1$, we have

$$\begin{aligned}\mathsf{E}(Y - \min(Y', 2))^2 &= \mathsf{E}(\min(Y, 2) - \min(Y', 2))^2 \\ &\leq 2\mathsf{E}|\min(Y, 2) - \min(Y', 2)| \\ &\leq 2\mathsf{E}|Y - Y'| \leq \frac{L}{\gamma^2} .\end{aligned}$$

The r.v. $A = Y' - \min(Y', 2)$ satisfies

$$A > 0 \Rightarrow A = Y' - 2 ,$$

so that if $t > 0$ we have $\mathsf{P}(A \geq t) = \mathsf{P}(Y' \geq t+2)$. Since Y' and a' have the same distribution, it holds:

$$\begin{aligned}\mathsf{E}(\min(Y',2) - Y')^2 = \mathsf{E}A^2 &= 2\int_0^\infty t\mathsf{P}(Y' \geq t+2)\mathrm{d}t \\ &= 2\int_0^\infty t\mathsf{P}(a' \geq t+2)\mathrm{d}t\,.\end{aligned}$$

To estimate $\mathsf{P}(a' \geq t)$, we write, for $\lambda > 0$

$$\begin{aligned}\mathsf{P}(a' \geq t) &\leq \exp(-\lambda\, t)\, \mathsf{E}\exp \lambda\, a' \\ &= \exp(-\lambda\, t)\big(\mathsf{E}\exp(\lambda \exp(-\gamma\, R\, X))\big)^R\end{aligned}$$

and, using (7.39) in the second inequality, and a power expansion of e^λ to obtain the third inequality, we get

$$\begin{aligned}\mathsf{E}\exp(\lambda \exp(-\gamma\, R\, X)) &= \sum_{p\geq 0} \frac{\lambda^p}{p!}\, \mathsf{E}\exp(-\gamma\, R\, p\, X) \\ &\leq 1 + \sum_{p\geq 1} \frac{\lambda^p}{p!\, p\, \gamma\, R} \leq 1 + \frac{e^\lambda}{\gamma\, R} \\ &\leq \exp\left(\frac{e^\lambda}{\gamma\, R}\right)\end{aligned}$$

so that

$$\mathsf{P}(a' \geq t) \leq \exp\left(\frac{e^\lambda}{\gamma} - \lambda\, t\right)\,.$$

Taking $\lambda = \log \gamma > 0$, we get

$$\mathsf{P}(a' \geq t) \leq L\, \gamma^{-t}$$

so that since $\gamma \geq 2$ we obtain

$$\int_0^\infty t\, \mathsf{P}(a' \geq t+2)\, \mathrm{d}t \leq \frac{L}{\gamma^2}\,. \qquad \square$$

Research Problem 7.6.10. (Level 2) Is it true that given an integer n, there exists a constant $K(\alpha, n)$, and independent r.v.s $U_1, \ldots, U_n$ of law μ with

$$\mathsf{E}\sum_{i\leq n}(u_{N,M}(i) - U_i)^2 \leq \frac{K(\alpha,n)}{N}\ ? \tag{7.98}$$

Proof of Theorem 7.1.2. We will stay somewhat informal in this proof. We write $A_{N,M} = \mathsf{E}\log Z_{N,M}$, so that

$$A_{N,M} - A_{N,M-1} = \mathsf{E}\log\frac{Z_{N,M}}{Z_{N,M-1}} = -\mathsf{E}\log u_{N,M}(M-1)$$

$$A_{N,M} - A_{N-1,M} = \mathsf{E}\log\frac{Z_{N,M}}{Z_{N-1,M}} = -\mathsf{E}\log w_{N,M}(N)\ .$$

By Theorem 7.1.2, these quantities have limits $-\int \log x\, \mathrm{d}\mu\,(x)$ and $-\int \log x\, \mathrm{d}\nu\,(x)$ respectively. (To obtain the required tightness, we observe that from (7.27), (7.28) and Markov's inequality we have $\mathsf{P}(u_{N,M}(M-1) < t) \leq Kt$ and $\mathsf{P}(w_{N,M}(N) < t) \leq Kt$.) Setting $M(R) = \lfloor R(1+\alpha)\rfloor$, we write

$$A_{N,M} - A_{1,1} = \mathrm{I} + \mathrm{II}\ ,$$

where

$$\mathrm{I} = \sum_{2\leq R\leq M} A_{R,M(R)} - A_{R-1,M(R)}$$

$$\mathrm{II} = \sum_{2\leq R\leq M} A_{R-1,M(R)} - A_{R-1,M(R-1)}\ .$$

For large R we have

$$A_{R,M(R)} - A_{R-1,M(R)} \simeq -\int \log x\, \mathrm{d}\nu(x)\ ,$$

and since $M(R) - 2 \leq M(R-1) \leq M(R) - 1$, we also have

$$A_{R-1,M(R)} - A_{R-1,M(R-1)} \simeq -(M(R) - M(R-1))\int \log x\, \mathrm{d}\mu(x)\ .$$

The result follows. □

A direction that should be pursued is the detailed study of Gibbs' measure; the principal difficulty might be to discover fruitful formulations. If G denotes Gibbs' measure, we should note the relation

$$G(\{\sigma(i) = j\}) = a(i,j)\frac{Z_{N,M}(i;j)}{Z_{N,M}} \simeq a(i,j)w_{N,M}(i)u_{N,M}(j)\ . \tag{7.99}$$

Also, if $i_1 \neq i_2$ and $j_1 \neq j_2$, we have

$$G(\{\sigma(i_1) = j_1\,;\, \sigma(i_2) = j_2\}) = a(i_1,i_2)a(j_1,j_2)\frac{Z_{N,M}(i_1,i_2;j_1,j_2)}{Z_{N,M}}\ . \tag{7.100}$$

One can generalize (7.7) to show that

$$\frac{Z_{N,M}(i_1,i_2;j_1,j_2)}{Z_{N,M}} \simeq w_{N,M}(i_1)w_{N,M}(i_2)u_{N,M}(j_1)u_{N,M}(j_2)$$

so comparing (7.99) and (7.100) we get

$$G(\{\sigma(i_1) = j_1\,;\, \sigma(i_2) = j_2\}) \simeq G(\{\sigma(i_1) = j_1\})G(\{\sigma(i_2) = j_2\})\ .$$

The problem however to find a nice formulation is that the previous relation holds for most values of j_1 and j_2 simply because both sides are nearly zero!

7.7 Notes and Comments

A recent paper [169] suggests that it could be of interest to investigate the following model. The configuration space consists of all pairs (A, σ) where A is a subset of $\{1, \ldots, N\}$, and where σ is a one to one map from A to $\{1, \ldots, N\}$. The Hamiltonian is then given by

$$H_N((A, \sigma)) = -C \operatorname{card} A + \beta N \sum_{i \in A} c(i, \sigma(i)), \tag{7.101}$$

where C is a constant and $c(i, j)$ are as previously. The idea of the Hamiltonian is that the term $-C \operatorname{card} A$ favors the pairs (A, σ) for which $\operatorname{card} A$ is large. It seems likely that, given C, results of the same nature as those we proved can be obtained for this model when $\beta \leq \beta(C)$, but that it will be difficult to prove the existence of a number β_0 such than these results hold for $\beta \leq \beta_0$, independently of the value of C, and even more difficult to prove that (as the results of [169] seem to indicate) they will hold for any value of C and of β.