Commun. Math. Phys. 325, 333–366 (2014)
Digital Object Identifier (DOI) 10.1007/s00220-013-1862-3

Communications in
Mathematical Physics

An Iterative Construction of Solutions of the TAP Equations for the Sherrington–Kirkpatrick Model

Erwin Bolthausen⋆

Institute of Mathematics, Universität Zürich, Zürich, Switzerland. E-mail: eb@math.uzh.ch

Received: 1 June 2012 / Accepted: 12 June 2013

Abstract: We propose an iterative scheme for the solutions of the TAP-equations in the Sherrington–Kirkpatrick model which is shown to converge up to and including the de Almeida–Thouless line. The main tool is a representation of the iterations which reveals an interesting structure of them. This representation does not depend on the temperature parameter, but for temperatures below the de Almeida–Thouless line, it contains a part which does not converge to zero in the limit.

1. Introduction

The TAP equations [9] for the Sherrington–Kirkpatrick model describe the quenched expectations of the spin variables in a large system.

The standard SK-model has the random Hamiltonian on $\Sigma_N \stackrel{\text{def}}{=} \{-1, 1\}^N,\ N \in \mathbb{N}$,

$$H_{N,\beta,h,\omega}(\boldsymbol{\sigma}) \stackrel{\text{def}}{=} -\beta \sum_{1\leq i<j\leq N} g_{ij}^{(N)}(\omega)\,\sigma_i\sigma_j - h\sum_{i=1}^{N}\sigma_i,$$

where $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_N) \in \Sigma_N$, $\beta > 0,\ h \geq 0$, and where the $g_{ij}^{(N)}$, $1 \leq i < j \leq N$, are i.i.d. centered Gaussian random variables with variance $1/N$, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We extend this matrix to a symmetric one, by putting $g_{ij} \stackrel{\text{def}}{=} g_{ji}$ for $i > j$, and $g_{ii} \stackrel{\text{def}}{=} 0$. The quenched Gibbs measure on Σ_N is

$$\frac{1}{Z_{N,\beta,h,\omega}} \exp\left[-H_{N,\beta,h,\omega}(\boldsymbol{\sigma})\right],$$

where $Z_{N,\beta,h,\omega} \stackrel{\text{def}}{=} \sum_{\boldsymbol{\sigma}} \exp\left[-H_{N,\beta,h,\omega}(\boldsymbol{\sigma})\right]$.

⋆Supported by an SNF grant No 200020-125247, and by the Humboldt Society.

We write $\langle \cdot \rangle_{N,\beta,h,\omega}$ for the expectation under this measure. We will often drop the indices N, β, h, ω if there is no danger of confusion. We set

$$m_i \stackrel{\text{def}}{=} \langle \sigma_i \rangle .$$

The TAP equations state that

$$m_i = \tanh\left(h + \beta \sum_{j=1}^{N} g_{ij} m_j - \beta^2 (1-q)\, m_i\right), \tag{1.1}$$

which have to be understood in a limiting sense, as $N \to \infty$. $q = q(\beta, h)$ is the solution of the equation

$$q = \int \tanh^2\left(h + \beta\sqrt{q}z\right) \phi\,(dz)\,, \tag{1.2}$$

where $\phi\,(dz)$ is the standard normal distribution. It is known that this equation has a unique solution $q > 0$ for $h > 0$ (see [7, Prop. 1.3.8]). If $h = 0$, then $q = 0$ is the unique solution if $\beta \leq 1$, and there are two other (symmetric) solutions when $\beta > 1$, which are supposed to be the relevant ones. Mathematically, the validity of the TAP equations has only been proved in the high temperature case, i.e. when β is small, although in the physics literature, it is claimed that they are valid also at low temperature, but there they have many solutions, and the Gibbs expectation has to be taken inside "pure states". For the best mathematical results, see [7, Chap. 1.7].

The appearance of the so-called Onsager term $\beta^2 (1-q)\, m_i$ is easy to understand. From standard mean-field theory, one would expect an equation

$$m_i = \tanh\left(h + \beta \sum_{j=1}^{N} g_{ij} m_j\right),$$

but one has to take into account the stochastic dependence between the random variables m_j and g_{ij}. In fact, it turns out that the above equation should be correct when one replaces m_j by $m_j^{(i)}$, where the latter is computed under a Gibbs average dropping the interactions with the spin i. Therefore $m_j^{(i)}$ is independent of the $g_{ik},\ 1 \leq k \leq N$, and one would get

$$m_i = \tanh\left(h + \beta \sum_{j=1}^{N} g_{ij} m_j^{(i)}\right). \tag{1.3}$$

The Onsager term is an Itô-type correction expanding the dependency of m_j on $g_{ji} = g_{ij}$, and replacing $m_j^{(i)}$ on the right hand side by m_j. The correction term is non-vanishing because of

$$\sum_j g_{ij}^2 \approx 1,$$

i.e. exactly for the same reason as in the Itô-correction in stochastic calculus. We omit the details which are explained in [5].

In the present paper, there are no results about SK itself. We introduce an iterative approximation scheme for solutions of the TAP equations which is shown to converge below and at the de Almeida–Thouless line, i.e. under condition (2.1) below (see [3]). This line is supposed to separate the high-temperature region from the low-temperature one, but although the full Parisi formula for the free energy of the SK-model has been proved by Talagrand [8], and with a different approach recently by Panchenko [6], there is no proof yet that the AT line is the correct phase separation line.

The scheme we propose is defined as follows: Consider sequences $\mathbf{m}^{(k)} = \left\{m_i^{(k)}\right\}_{1\leq i\leq N}$, $k \in \mathbb{N}$, of random variables by defining recursively

$$\mathbf{m}^{(0)} \stackrel{\text{def}}{=} 0,\ \mathbf{m}^{(1)} \stackrel{\text{def}}{=} \sqrt{q}\mathbf{1}.$$

$\mathbf{1}$ here the vector with coordinates all 1, and $q = q(\beta, h)$ is the unique solution of (1.2). Then, defining tanh componentwise, we set

$$\mathbf{m}^{(k+1)} \stackrel{\text{def}}{=} \tanh\left(h + \beta\,\mathbf{g}\mathbf{m}^{(k)} - \beta^2(1-q)\,\mathbf{m}^{(k-1)}\right), \quad k \geq 1.$$

This iterative scheme reveals, we believe, an interesting structure of the dependence of the m_i on the family $\{g_{ij}\}$, even below the AT line. The main technical result, Proposition 2.5 is proved at all temperatures, but beyond the AT-line, it does not give much information.

It may be useful to sketch the first two steps: $m_i^{(1)} = \sqrt{q}$ and $m_i^{(2)} = \tanh\left(h + \beta\sqrt{q}\xi_i^{(1)}\right)$, where $\xi_i^{(1)} \stackrel{\text{def}}{=} \sum_j g_{ij}$. Then by the law of large numbers,

$$\frac{1}{N}\sum_{i=1}^{N} m_i^{(2)} \simeq \int \tanh\left(h + \beta\sqrt{q}z\right)\phi(dz) \stackrel{\text{def}}{=} \gamma_1,$$

and

$$\frac{1}{N}\sum_{i=1}^{N} m_i^{(2)2} \simeq \int \tanh^2\left(h + \beta\sqrt{q}z\right)\phi(dz) = q.$$

The next step is more interesting, as there the Onsager correction appears:

$$m_i^{(3)} = \tanh\left(h + \beta\sum_j g_{ij}m_j^{(2)} - \beta^2(1-q)\sqrt{q}\right).$$

In order to be able to compute inner products $N^{-1}\sum_{k=1}^{N} m_k^{(i)}m_k^{(j)}$, we replace $\mathbf{g}$ by a matrix $\mathbf{g}^{(2)}$ which is independent of $\mathbf{m}^{(2)}$. As this latter vector depends on the g_{ij} only through the $\xi_i^{(1)}$, we can obtain this by a linear procedure. The exact formula is somewhat complicated, but the leading correction is easily described: The shift from g to $g^{(2)}$ which is independent of $\xi^{(1)}$ essentially is:

$$g_{ij}^{(2)} \simeq g_{ij} - N^{-1}\left(\xi_i^{(1)} + \xi_j^{(1)}\right).$$

By the law of large numbers, this leads to a correction inside tanh of $m_i^{(3)}$:

$$\beta N^{-1}\sum_j\left(\xi_i^{(1)}+\xi_j^{(1)}\right)m_j^{(2)}\simeq\beta\gamma_1\xi_i^{(1)}+\beta\int x\tanh\left(h+\beta\sqrt{q}x\right)\phi\left(dx\right)$$
$$=\beta\gamma_1\xi_i^{(1)}+\beta^2\sqrt{q}\left(1-q\right),$$

which implies that

$$m_i^{(3)}\simeq\tanh\left(h+\beta\sum_j g_{ij}^{(2)}m_j^{(2)}+\beta\gamma_1\xi_i^{(1)}\right).$$

Using now the fact that $\mathbf{g}^{(2)}$ is independent of the $m_i^{(2)}$, $\xi_i^{(1)}$, one can easily evaluate inner products like $N^{-1}\sum_{i=1}^N m_i^{(2)}m_i^{(3)}$ in the $N\to\infty$ limit.

More interesting features appear in the next iteration

$$m_i^{(4)}=\tanh\left(h+\beta\sum_j g_{ij}m_j^{(3)}-\beta^2\left(1-q\right)m_i^{(2)}\right).$$

As the $m_i^{(3)}$ no longer depend linearly on $\mathbf{g}$, one first does the replacement from $\mathbf{g}$ to $\mathbf{g}^{(2)}$, and then one chooses $\mathbf{g}^{(3)}$, *conditionally on* $\xi^{(1)}$, independent of the $\sum_j g_{ij}^{(2)}m_j^{(2)}$. Fixing $\xi^{(1)}$, the $m_j^{(2)}$ are constant, and therefore, conditionally on $\xi^{(1)}$, $\mathbf{m}^{(3)}$ depends linearly on $\mathbf{g}^{(2)}$ and one obtains $\mathbf{g}^{(3)}$ by a "conditionally linear" transformation. After doing a similar application of the LLN as above, the replacements lead to

$$m_i^{(4)}\simeq\tanh\left(h+\beta\sum_j g_{ij}^{(3)}m_j^{(3)}+\beta\gamma_1\xi_i^{(1)}+\beta\gamma_2\xi_i^{(2)}\right),$$

with

$$\xi_i^{(2)}\overset{\text{def}}{=}\sum_j g_{ij}^{(2)}\frac{m_j^{(2)}-\gamma_1}{\sqrt{q-\gamma_1}},$$

and some $\gamma_2>0$, and again, in this form, it is not difficult to evaluate $N^{-1}\sum_{k=1}^N m_k^{(i)}m_k^{(j)}$ for $i,j\leq 4$ and $N\to\infty$, by applying the law of large numbers, but in a conditional version. The details of this are explained later.

In this way one can go on, and the outcome is a representation

$$m_i^{(k)}\simeq\tanh\left(h+\beta\sum_j g_{ij}^{(k-1)}m_j^{(k-1)}+\beta\sum_{r=1}^{k-2}\gamma_r\xi_i^{(r)}\right). \tag{1.4}$$

The interesting structure is that at every step of the iteration, the additional dependency between $\mathbf{g}$ and $\mathbf{m}^{(k-1)}$ is shifted into an additional term $\gamma_{k-2}\xi^{(k-2)}$ which is tractable. We then prove that β is below or at the AT-line if and only if the first part $\sum_j g_{ij}^{(k-1)}m_j^{(k-1)}$ asymptotically vanishes as $k\to\infty$ (in a way to be made precise) which leads to the convergence of the iterative procedure.

Our method of sequential conditioning used to prove Proposition 2.5 has been used in other contexts by Bayati and Montanari [1] and Bayati et al. [2].

We finish the section by introducing some notations.

If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$, we write

$$\langle \mathbf{x}, \mathbf{y} \rangle \stackrel{\text{def}}{=} \frac{1}{N} \sum_{i=1}^{N} x_i y_i, \quad \|\mathbf{x}\| \stackrel{\text{def}}{=} \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}.$$

As mentioned above, we suppress N in notations as far as possible, but this parameter is present everywhere.

We also define the $N \times N$-matrix $\mathbf{x} \otimes_{\mathrm{s}} \mathbf{y}$, $\mathbf{x} \otimes \mathbf{y}$, by

$$(\mathbf{x} \otimes_{\mathrm{s}} \mathbf{y})_{i,j} \stackrel{\text{def}}{=} \frac{1}{N} \left(x_i y_j + x_j y_i \right), \quad \mathbf{x} \otimes \mathbf{y} \stackrel{\text{def}}{=} \frac{x_i y_j}{N}. \tag{1.5}$$

If $\mathbf{A}$ is an $N \times N$-matrix and $\mathbf{x} \in \mathbb{R}^N$, the vector $\mathbf{Ax}$ is defined in the usual way (interpreting vectors in $\mathbb{R}^N$ as column matrices). If $f : \mathbb{R} \to \mathbb{R}$ is a function and $\mathbf{x} \in \mathbb{R}^N$ we simply write $f(\mathbf{x})$ for the vector obtained by applying f to the coordinates.

$\mathbf{g} = \left(g_{ij} \right)$ is a Gaussian $N \times N$-matrix where the g_{ij} for $i < j$ are independent centered Gaussians with variance $1/N$, and where $g_{ij} = g_{ji}$, $g_{ii} = 0$. We will exclusively reserve the notation $\mathbf{g}$ for such a Gaussian matrix.

We will use $Z, Z', Z_1, Z_2, \ldots$ as generic standard Gaussians. Whenever several of them appear in the same formula, they are assumed to be independent, without special mentioning. We then write E when taking expectations with respect to them. (This notation is simply an outflow of the abhorrence probabilists have of using integral signs, as John Westwater once put it).

If $\{X_N\}$, $\{Y_N\}$ are two sequences of real random variables, defined on $(\Omega, \mathcal{F}, \mathbb{P})$, we write

$$X_N \simeq Y_N,$$

provided there exists a constant $C > 0$ such that

$$\mathbb{P}\left(|X_N - Y_N| \geq t\right) \leq C \exp\left[-t^2 N / C\right]$$

for all $N \in \mathbb{N}$, $0 < t \leq 1$.

Clearly, if $X_N \simeq Y_N$, and $X'_N \simeq Y'_N$, then $X_N + X'_N \simeq Y_N + Y'_N$.

If $\mathbf{X}^{(N)} = \left(X_i^{(N)} \right)_{i \leq N}$, $\mathbf{Y}^{(N)} = \left(Y_i^{(N)} \right)_{i \leq N}$ are two sequences of random vectors in $\mathbb{R}^N$, we write $\mathbf{X}^{(N)} \approx \mathbf{Y}^{(N)}$ if

$$\frac{1}{N} \sum_{i=1}^{N} \left| X_i^{(N)} - Y_i^{(N)} \right| \simeq 0.$$

We will use $C > 0$ as a generic positive constant, not necessarily the same at different occurrences. It may depend on β, h, and on the level k of the approximation scheme appearing in the next section, but on nothing else, unless stated otherwise.

In order to avoid endless repetitions of the parameters h and β, we use the abbreviation

$$\mathrm{Th}(x) = \tanh(h + \beta x).$$

We always assume $h \neq 0$, and as there is a symmetry between the signs, we assume $h > 0$. $q = q(\beta, h)$ will exclusively be used for the unique solution of (1.2). In the case $h = 0,\ \beta > 1$, there is a unique solution of (1.2) which is positive. Proposition 2.5 is valid in this case, too, but this does not lead to a useful result. So, we stick to the $h > 0$ case.

Gaussian random variables are always assumed to be centered.

2. The Recursive Scheme for the Solutions of the TAP Equations

k will exclusively be used to number the level of the iteration. Our main result is

Theorem 2.1. *Assume $h > 0$. If $\beta > 0$ is below the AT-line, i.e. if*

$$\beta^2 E \cosh^{-4}\left(h + \beta\sqrt{q}Z\right) \leq 1, \tag{2.1}$$

then

$$\lim_{k,k'\to\infty} \limsup_{N\to\infty} \mathbb{E}\left\|\mathbf{m}^{(k)} - \mathbf{m}^{(k')}\right\|^2 = 0.$$

If there is strict inequality in (2.1), then there exist $0 < \lambda(\beta, h) < 1$, and $C > 0$, such that for all k,

$$\limsup_{N\to\infty} \mathbb{E}\left\|\mathbf{m}^{(k+1)} - \mathbf{m}^{(k)}\right\|^2 \leq C\lambda^k.$$

The theorem is a straightforward consequence of a computation of the inner products $\langle\mathbf{m}^{(i)}, \mathbf{m}^{(j)}\rangle$. We explain that first. The actual computation of these inner products will be quite involved and will depend on clarifying the structural dependence of $\mathbf{m}^{(k)}$ on $\mathbf{g}$.

As we assume $h > 0$, we have $q > 0$. We define a function $\psi : [0, q] \to \mathbb{R}$ by

$$\psi(t) \stackrel{\text{def}}{=} E\,\text{Th}\left(\sqrt{t}Z + \sqrt{q-t}Z'\right)\text{Th}\left(\sqrt{t}Z + \sqrt{q-t}Z''\right),$$

where Z, Z', Z'', as usual, are independent standard Gaussians. Remember that $\text{Th}(x) = \tanh(h + \beta x)$.

Let $\alpha \stackrel{\text{def}}{=} E\,\text{Th}\left(\sqrt{q}Z\right) > 0$.

Lemma 2.2. (a) *ψ satisfies $0 < \psi(0) = \alpha^2 < \psi(q) = q$, and is strictly increasing and convex on $[0, q]$.*

(b)

$$\psi'(q) = \beta^2 E \cosh^{-4}\left(h + \beta\sqrt{q}Z\right).$$

Proof. $\psi(0) = \alpha^2$, and $\psi(q) = q$ are evident by the definition of α, q. We compute the first two derivatives of ψ :

$$\psi'(t) = \frac{1}{\sqrt{t}} E\Big[Z\,\text{Th}'\left(\sqrt{t}Z + \sqrt{q-t}Z'\right)\text{Th}\left(\sqrt{t}Z + \sqrt{q-t}Z''\right)\Big]$$
$$-\frac{1}{\sqrt{q-t}} E\Big[Z'\,\text{Th}'\left(\sqrt{t}Z + \sqrt{q-t}Z'\right)\text{Th}\left(\sqrt{t}Z + \sqrt{q-t}Z''\right)\Big]$$

$$\begin{aligned}
&= E\,\mathrm{Th}''\left(\sqrt{t}Z+\sqrt{q-t}Z'\right)\mathrm{Th}\left(\sqrt{t}Z+\sqrt{q-t}Z''\right)\\
&\quad+E\,\mathrm{Th}'\left(\sqrt{t}Z+\sqrt{q-t}Z'\right)\mathrm{Th}'\left(\sqrt{t}Z+\sqrt{q-t}Z''\right)\\
&\quad-E\,\mathrm{Th}''\left(\sqrt{t}Z+\sqrt{q-t}Z'\right)\mathrm{Th}\left(\sqrt{t}Z+\sqrt{q-t}Z''\right)\\
&= E\,\mathrm{Th}'\left(\sqrt{t}Z+\sqrt{q-t}Z'\right)\mathrm{Th}'\left(\sqrt{t}Z+\sqrt{q-t}Z''\right),
\end{aligned}$$

the second equality by Gaussian partial integration.

Differentiating once more, we get

$$\psi''(t) = E\left(\mathrm{Th}''\left(\sqrt{t}Z+\sqrt{q-t}Z'\right)\mathrm{Th}''\left(\sqrt{t}Z+\sqrt{q-t}Z''\right)\right).$$

In both expressions, we can first integrate out Z', Z'', getting

$$\psi'(t) = \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}\mathrm{Th}'\left(\sqrt{t}x+\sqrt{q-t}y\right)\phi(dy)\right]^2\phi(dx) > 0,$$

and the similar expression for ψ'' with Th' replaced by Th''. So, we see that ψ is increasing and convex. Furthermore, as

$$\begin{aligned}
\mathrm{Th}'(x) &= \beta\tanh'(\beta x+h) = \beta\left(1-\tanh^2(\beta x+h)\right)\\
&= \frac{\beta}{\cosh^2(\beta x+h)},
\end{aligned}$$

we get

$$\psi'(q) = E\,\mathrm{Th}'\left(\sqrt{q}Z\right)^2 = \beta^2 E\cosh^{-4}\left(h+\beta\sqrt{q}Z\right).$$

□

Corollary 2.3. *If* (2.1) *is satisfied, then* q *is the only fixed point of* ψ *in the interval* $[0, q]$. *If* (2.1) *is not satisfied then there is a unique fixed point of* $\psi(t) = t$ *inside the interval* $(0, q)$.

We define sequences $\{\rho_k\}_{k\geq 1}$, $\{\gamma_k\}_{k\geq 1}$ recursively by $\gamma_1 \stackrel{\text{def}}{=} \alpha$, $\rho_1 \stackrel{\text{def}}{=} \gamma_1\sqrt{q}$, and for $k \geq 2$,

$$\begin{aligned}
\rho_k &\stackrel{\text{def}}{=} \psi(\rho_{k-1}),\\
\gamma_k &\stackrel{\text{def}}{=} \frac{\rho_k - \Gamma_{k-1}^2}{\sqrt{q-\Gamma_{k-1}^2}},
\end{aligned}$$

where

$$\Gamma_m^2 \stackrel{\text{def}}{=} \sum_{j=1}^m \gamma_j^2, \quad \Gamma_0^2 \stackrel{\text{def}}{=} 0.$$

In order for γ_k to be well defined, we have to prove recursively that $\Gamma_{k-1}^2 < q$.

Lemma 2.4. (a) *For all* $k \in \mathbb{N}$,

$$\Gamma_{k-1}^2 < \rho_k < q.$$

(b) *If* (2.1) *is satisfied, then*

$$\lim_{k\to\infty} \rho_k = q, \quad \lim_{k\to\infty} \Gamma_k^2 = q.$$

(c) *If there is strict inequality in* (2.1) *, then* Γ_k^2 *and* ρ_k *converge to* q *exponentially fast.*

Proof. (a) $\rho_k < q$ for all k is evident.
We prove by induction on k that $\rho_k > \Gamma_{k-1}^2$. For $k = 1$, as $\rho_1 = \gamma_1\sqrt{q}$, the statement follows.
Assume that it is true for k. Then

$$\gamma_k = \frac{\rho_k - \Gamma_{k-1}^2}{\sqrt{q - \Gamma_{k-1}^2}} < \sqrt{\rho_k - \Gamma_{k-1}^2},$$

i.e. $\rho_k > \Gamma_k^2$. As $\rho_{k+1} > \rho_k$, the statement follows.

(b) Evidently $\lim_{k\to\infty} \rho_k = q$ if (2.1) is satisfied. The sequence $\{\Gamma_k^2\}$ is increasing and bounded (by q). If $\zeta \stackrel{\text{def}}{=} \lim_{k\to\infty} \Gamma_k^2 < q$, then $\lim_{k\to\infty} \gamma_k = \sqrt{q-\zeta} > 0$, a contradiction to the boundedness of $\{\Gamma_k^2\}$.

(c) Linearization of ψ around q easily shows that the convergence is exponentially fast if $\psi'(q) < 1$. □

Remark that by (a) of the above lemma, one has $\gamma_k > 0$ for all k.

Let Π_j be the orthogonal projection in $\mathbb{R}^N$, with respect to the inner product $\langle\cdot,\cdot\rangle$, onto span $\left(\mathbf{m}^{(1)}, \ldots, \mathbf{m}^{(j)}\right)$. Remember that the inner product is N^{-1} times the standard inner product in $\mathbb{R}^N$. We set

$$\mathbf{M}^{(k,j)} \stackrel{\text{def}}{=} \mathbf{m}^{(k)} - \Pi_j\left(\mathbf{m}^{(k)}\right), \quad j < k, \tag{2.2}$$

and

$$\mathbf{M}^{(k)} \stackrel{\text{def}}{=} \mathbf{M}^{(k,k-1)}. \tag{2.3}$$

Let

$$\phi^{(k)} \stackrel{\text{def}}{=} \frac{\mathbf{M}^{(k)}}{\left\|\mathbf{M}^{(k)}\right\|} \tag{2.4}$$

if $\left\|\mathbf{M}^{(k)}\right\| \neq 0$. In case $\mathbf{m}^{(k)} \in$ span $\left(\mathbf{m}^{(1)}, \ldots, \mathbf{m}^{(k-1)}\right)$, we define $\phi^{(k)} \stackrel{\text{def}}{=} \mathbf{1}$, to have it defined everywhere, but we will see that this happens only with exponentially small probability. Remark that $\phi^{(1)} = \mathbf{1}$.

The key result is:

Proposition 2.5. *For all* $k \in \mathbb{N}$

$$\left\|\mathbf{m}^{(k)}\right\|^2 \simeq q, \tag{2.5}$$

and for $1 \leq j < k$,

$$\left\langle \mathbf{m}^{(j)}, \mathbf{m}^{(k)} \right\rangle \simeq \rho_j, \tag{2.6}$$

$$\left\langle \phi^{(j)}, \mathbf{m}^{(k)} \right\rangle \simeq \gamma_j. \tag{2.7}$$

Proof of Theorem 2.1 from Proposition 2.5. As the variables $\mathbf{m}^{(k)}$ are bounded, (2.5) implies

$$\lim_{N\to\infty} \mathbb{E} \left\| \mathbf{m}^{(k)} \right\|^2 = q,$$

and for $j < k$,

$$\lim_{N\to\infty} \mathbb{E} \left\langle \mathbf{m}^{(j)}, \mathbf{m}^{(k)} \right\rangle = \rho_j.$$

Therefore, for $k' < k$,

$$\mathbb{E} \left\| \mathbf{m}^{(k')} - \mathbf{m}^{(k)} \right\|^2 = \mathbb{E} \left\| \mathbf{m}^{(k)} \right\|^2 + \mathbb{E} \left\| \mathbf{m}^{(k')} \right\|^2 - 2\mathbb{E} \left\langle \mathbf{m}^{(k)}, \mathbf{m}^{(k')} \right\rangle.$$

Taking the $N \to \infty$, this converges to $2q - 2\rho_{k'}$. From Lemma 2.4, the claim follows. □

Remark 2.6. Proposition 2.5 is true for all temperatures. However, beyond the AT-line, it does not give much information on the behavior of the $\mathbf{m}^{(k)}$ for large k. If (2.1) is not satisfied, then there exists a unique number $q^* \in (0, q)$ with $\psi(q^*) = q^*$, and from $\psi', \psi'' > 0$ it easily follows that $\lim_{k\to\infty} \rho_k = q^*$. Therefore, it follows from Proposition 2.5 that

$$\lim_{k,k'\to\infty} \lim_{N\to\infty} \mathbb{E} \left\| \mathbf{m}^{(k')} - \mathbf{m}^{(k)} \right\|^2 = 2\left(q - q^*\right) > 0.$$

The main task is to prove Proposition 2.5. It follows by an involved induction argument.

Lemma 2.7. (2.7) *is a consequence of* (2.5) *and* (2.6).

Proof. We do induction on j.

For $j = 1$, we have $\phi^{(1)} = \mathbf{1}$, $\mathbf{m}^{(1)} = \sqrt{q}\mathbf{1}$, and therefore

$$\left\langle \phi^{(1)}, \mathbf{m}^{(k)} \right\rangle = \frac{1}{\sqrt{q}} \left\langle \mathbf{m}^{(1)}, \mathbf{m}^{(k)} \right\rangle \simeq \frac{\rho_1}{\sqrt{q}} = \gamma_1.$$

Let $j \geq 2$. Then $\phi^{(j)} = \mathbf{M}^{(j)} / \left\| \mathbf{M}^{(j)} \right\|$, and

$$\left\| \mathbf{M}^{(j)} \right\|^2 = \left\| \mathbf{m}^{(j)} \right\|^2 - \sum_{s=1}^{j-1} \left\langle \mathbf{m}^{(j)}, \phi^{(s)} \right\rangle^2.$$

As $\left|\left\langle \mathbf{m}^{(j)}, \phi^{(s)} \right\rangle\right|$ is bounded, it follows from the induction hypothesis that with

$$\delta_j \stackrel{\text{def}}{=} \sqrt{q - \Gamma_{j-1}^2/2} > 0, \tag{2.8}$$

one has

$$\mathbb{P}\left(\left\|\mathbf{M}^{(j)}\right\| \leq \delta_j\right) \leq C \exp\left[-N/C\right],$$

and that

$$\left\|\mathbf{M}^{(j)}\right\| \simeq \sqrt{q - \Gamma_{j-1}^2}. \tag{2.9}$$

If now $k > j$, then

$$\left\langle \mathbf{m}^{(k)}, \phi^{(j)} \right\rangle = \frac{\left\langle \mathbf{m}^{(k)}, \mathbf{m}^{(j)} \right\rangle - \sum_{s=1}^{j-1} \left\langle \mathbf{m}^{(k)}, \phi^{(s)} \right\rangle \left\langle \mathbf{m}^{(j)}, \phi^{(s)} \right\rangle}{\left\|\mathbf{M}^{(j)}\right\|},$$

and using (2.9), (2.6), and the induction hypothesis, we get

$$\left\langle \mathbf{m}^{(k)}, \phi^{(j)} \right\rangle \simeq \frac{\rho_j - \Gamma_{j-1}^2}{\sqrt{q - \Gamma_{j-1}^2}} = \gamma_j.$$

□

Definition 2.8. *If* $J \in \mathbb{N}$*, we say that* COND (J) *holds if* (2.5) *and* (2.6) *are true for* $k \leq J$.

COND (1) is trivially true.

Remark 2.9. Assume COND (J), and take δ_j as in (2.8),

$$A_J \overset{\text{def}}{=} \bigcap_{j=1}^{J} \left\{ \left\|\mathbf{M}^{(j)}\right\| > \delta_j \right\}, \tag{2.10}$$

which then satisfies

$$\mathbb{P}\left(A_J\right) \geq 1 - C_J \exp\left[-N/C_J\right]. \tag{2.11}$$

Furthermore, there exist constants $c_J > 0$, depending only on J, which can be expressed in terms of $\delta_j,\ j \leq J$, such that $\left|\phi_i^{(k)}(\omega)\right| \leq c_J$ for $\omega \in A_J$, and all $k \leq J$, all N, and all $i \leq N$. This is easily seen from (2.4) and $\left|m_i^{(k)}(\omega)\right| \leq 1$.

The rest of the paper is the proof that

$$\text{COND}\,(J) \Longrightarrow \text{COND}\,(J+1)\,. \tag{2.12}$$

In the course of the proof, we find the alternative representation of the $\mathbf{m}^{(k)}$ (1.4), precisely formulated in (5.13).

3. Iterative Modifications of the Interaction Variables

Let $\mathcal{G}$ be a sub-σ-field of $\mathcal{F}$, and $\mathbf{y} = \left(y_{ij}\right)_{1\leq i,j\leq N}$ be a random matrix. We are only interested in the case where $\mathbf{y}$ is symmetric and 0 on the diagonal, but this is not important for the moment. We assume that $\mathbf{y}$ is jointly Gaussian, conditioned on $\mathcal{G}$, i.e. there is a positive semidefinite, $\mathcal{G}$-m.b. $N^2 \times N^2$-matrix Γ such that

$$E\left(\exp\left[i\sum_{k,j} t_{kj}y_{kj}\right]\middle|\mathcal{G}\right) = \exp\left[-\frac{1}{2}\sum_{k,k',j,j'} t_{kj}\Gamma_{kj,k'j'}t_{k'j'}\right].$$

(We do not assume that $\mathbf{y}$ is Gaussian, unconditionally.) Consider a $\mathcal{G}$-measurable random vector $\mathbf{x}$, and the linear space of random variables

$$\mathcal{L} \stackrel{\text{def}}{=} \left\{\sum_{i=1}^{N} a_i\,(\mathbf{yx})_i : a_1,\ldots,a_N\ \mathcal{G}-\text{measurable}\right\}.$$

We consider the linear projection $\pi_{\mathcal{L}}(\mathbf{y})$ of $\mathbf{y}$ onto $\mathcal{L}$, which is defined to be the unique matrix with components $\pi_{\mathcal{L}}(y_{ij})$ in $\mathcal{L}$ which satisfy

$$\mathbb{E}\left(\left\{y_{ij} - \pi_{\mathcal{L}}(y_{ij})\right\}U|\mathcal{G}\right) = 0,\ \forall U \in \mathcal{L}.$$

As $\mathbf{y}$ is assumed to be conditionally Gaussian, given $\mathcal{G}$, it follows that $\mathbf{y} - \pi_{\mathcal{L}}(\mathbf{y})$ is conditionally independent of the variables in $\mathcal{L}$, given $\mathcal{G}$.

If $\mathbf{y}$ is symmetric, then clearly $\pi_{\mathcal{L}}(\mathbf{y})$ is symmetric, too.

Remark 3.1. If κ is a $\mathcal{G}$-measurable real-valued random variable then $\kappa\mathbf{y}$ is conditionally Gaussian as well and

$$\kappa\pi_{\mathcal{L}}(\mathbf{y}) = \pi_{\mathcal{L}}(\kappa\mathbf{y}).$$

Remark also that

$$(\mathbf{y} - \pi_{\mathcal{L}}(\mathbf{y}))\,\mathbf{x} = \mathbf{yx} - \pi_{\mathcal{L}}(\mathbf{yx}) = 0, \tag{3.1}$$

as $\mathbf{yx} \in \mathcal{L}$.

Using this construction, we define a sequence $\mathbf{g}^{(k)}$, $k \geq 1$ of matrices, and a sequence $\{\mathcal{F}_k\}$ of sub-σ-fields of $\mathcal{F}$, starting with $\mathbf{g}^{(1)} \stackrel{\text{def}}{=} \mathbf{g}$, and $\mathcal{F}_{-1} = \mathcal{F}_0 \stackrel{\text{def}}{=} \mathcal{N}$, the set of $\mathbb{P}$-null sets. The construction is done in such a way that

(C1) $\mathbf{g}^{(k)}$ is conditionally Gaussian, given $\mathcal{F}_{k-1}$.
(C2) $\mathbf{m}^{(k)}$, $\mathbf{M}^{(k)}$, and $\phi^{(k)}$ are $\mathcal{F}_{k-1}$-measurable.

Using that we define

$$\mathbf{g}^{(k+1)} \stackrel{\text{def}}{=} \mathbf{g}^{(k)} - \pi_{\mathcal{L}_k}\left(\mathbf{g}^{(k)}\right), \tag{3.2}$$

with

$$\mathcal{L}_k \stackrel{\text{def}}{=} \left\{\sum_{i=1}^{N} a_i\left(\mathbf{g}^{(k)}\mathbf{M}^{(k)}\right)_i : a_i\ \mathcal{F}_{k-1}\text{-measurable}\right\},$$

i.e. we perform the above construction with $\mathcal{G} = \mathcal{F}_{k-1}$ and $\mathbf{x} = \mathbf{M}^{(k)}$.

Furthermore, we define

$$\mathcal{F}_{k+1} \stackrel{\text{def}}{=} \sigma\left(\mathcal{F}_k, \xi^{(k+1)}\right),$$

where

$$\xi^{(k)} \stackrel{\text{def}}{=} \mathbf{g}^{(k)}\phi^{(k)}.$$

There is a "nicer" representation of $\mathcal{F}_k$ stated below in Remark 3.4, but the one given above is more convenient for the moment.

In order that the construction is well defined, we have to inductively prove the properties **(C1)** and **(C2).** We actually prove a condition which is stronger than **(C1)**:

(C1') Conditionally on $\mathcal{F}_{k-2}$, $\mathbf{g}^{(k)}$ is Gaussian, and conditionally independent of $\mathcal{F}_{k-1}$.

(C1') implies that $\mathbf{g}^{(k)}$ is conditionally Gaussian, given $\mathcal{F}_{k-1}$, and the conditional law, given $\mathcal{F}_{k-1}$, is the same as given $\mathcal{F}_{k-2}$.

Inductive proof of (C1') and (C2). The case $k = 1$ is trivial. We first prove (C2) for $k \geq 2$, using (C1'), (C2) up to $k-1$. We claim that

$$\mathbf{m}^{(k)} = \mathrm{Th}\left(\mathbf{g}^{(k-1)}\mathbf{M}^{(k-1)} + \mathbf{R}^{(k-2)}\right), \tag{3.3}$$

where $\mathbf{R}^{(k-2)}$ stands for a generic $\mathcal{F}_{k-2}$-measurable random variable, not necessarily the same at different occurrences.

As $\mathbf{g}^{(k-1)}\mathbf{M}^{(k-1)} = \left\|\mathbf{M}^{(k-1)}\right\| \xi^{(k-1)}$, and $\mathbf{M}^{(k-1)}$ is $\mathcal{F}_{k-2}$-measurable, by the induction hypothesis, it follows from (3.3) that $\mathbf{m}^{(k)}$ is $\mathcal{F}_{k-1}$-measurable The statements for $\mathbf{M}^{(k)}$, $\phi^{(k)}$ are then trivial consequences.

We therefore have to prove (3.3). We prove by induction on j that

$$\mathbf{m}^{(k)} = \mathrm{Th}\left(\mathbf{g}^{(j)}\mathbf{M}^{(k-1,j-1)} + \mathbf{R}^{(k-2)}\right). \tag{3.4}$$

The case $j = 1$ follows from the definition of $\mathbf{m}^{(k)}$, and the case $j = k-1$ is (3.3).

Assume that (3.4) is true for $j < k-1$. We replace $\mathbf{g}^{(j)}$ by $\mathbf{g}^{(j+1)}$ through the recursive definition

$$\begin{aligned}\mathbf{m}^{(k)} &= \mathrm{Th}\left(\mathbf{g}^{(j+1)}\mathbf{M}^{(k-1,j-1)} + \pi_{\mathcal{L}_j}\left(\mathbf{g}^{(j)}\right)\mathbf{M}^{(k-1,j-1)} + \mathbf{R}^{(k-2)}\right)\\ &= \mathrm{Th}\left(\mathbf{g}^{(j+1)}\mathbf{M}^{(k-1,j-1)} + \mathbf{R}^{(k-2)}\right),\end{aligned}$$

as $\pi_{\mathcal{L}_j}\left(\mathbf{g}^{(j)}\right)$ is $\mathcal{F}_j$-measurable and therefore $\pi_{\mathcal{L}_j}\left(\mathbf{g}^{(j)}\right)\mathbf{M}^{(k-1,j-1)}$ is $\mathcal{F}_{k-2}$-measurable.

Using (3.1), one gets $\mathbf{g}^{(j+1)}\mathbf{M}^{(j)} = 0$, and therefore

$$\mathbf{g}^{(j+1)}\mathbf{M}^{(k-1,j-1)} = \mathbf{g}^{(j+1)}\mathbf{M}^{(k-1,j)}.$$

This proves (3.3), and therefore (C2) for k.

We next prove (C1') for k,

$$\mathbf{g}^{(k)} \stackrel{\text{def}}{=} \mathbf{g}^{(k-1)} - \pi_{\mathcal{L}_{k-1}}\left(\mathbf{g}^{(k-1)}\right).$$

We condition on $\mathcal{F}_{k-2}$. By (C2), $\mathbf{M}^{(k-1)}$ is $\mathcal{F}_{k-2}$-measurable. As $\mathbf{g}^{(k-1)}$, conditioned on $\mathcal{F}_{k-3}$, is Gaussian, and independent of $\mathcal{F}_{k-2}$, it has the same distribution also conditioned on $\mathcal{F}_{k-2}$. By the construction of $\mathbf{g}^{(k)}$, this variable is conditioned on $\mathcal{F}_{k-2}$, independent of $\mathcal{F}_{k-1}$, and conditionally Gaussian. □

Lemma 3.2. *For $m < k$ one has*

$$\mathbf{g}^{(k)}\phi^{(m)} = 0.$$

Proof. The proof is by induction on k. For $k = 1$, there is nothing to prove.

Assume that the statement is proved up to k. We want to prove $\mathbf{g}^{(k+1)}\phi^{(m)} = 0$ for $m \leq k$. The case $m = k$ is covered by (3.1). For $m < k$, it follows by Remark 3.1, as $\phi^{(m)}$ is $\mathcal{F}_{k-1}$-measurable, that

$$\pi_{\mathcal{L}_k}\left(\mathbf{g}^{(k)}\right)\phi^{(m)} = \pi_{\mathcal{L}_k}\left(\mathbf{g}^{(k)}\phi^{(m)}\right),$$

and therefore

$$\begin{aligned}\mathbf{g}^{(k+1)}\phi^{(m)} &= \mathbf{g}^{(k)}\phi^{(m)} - \pi_{\mathcal{L}_k}\left(\mathbf{g}^{(k)}\right)\phi^{(m)} \\ &= \mathbf{g}^{(k)}\phi^{(m)} - \pi_{\mathcal{L}_k}\left(\mathbf{g}^{(k)}\phi^{(m)}\right) = 0,\end{aligned}$$

as $\mathbf{g}^{(k)}\phi^{(m)} = 0$ by the induction hypothesis. □

Lemma 3.3. *If $m < k$, then*

$$\sum_i \xi_i^{(k)}\phi_i^{(m)} = 0.$$

Proof.

$$\begin{aligned}\sum_i \xi_i^{(k)}\phi_i^{(m)} &= \sum_i\sum_j g_{ij}^{(k)}\phi_j^{(k)}\phi_i^{(m)} \\ &= \sum_j \phi_j^{(k)}\sum_i g_{ij}^{(k)}\phi_i^{(m)} \\ &= \sum_j \phi_j^{(k)}\sum_i g_{ji}^{(k)}\phi_i^{(m)} = 0\end{aligned}$$

for $m < k$, by the symmetry of $\mathbf{g}^{(k)}$ and the previous lemma. □

Remark 3.4. $\mathcal{F}_k$ has a more straightforward description in terms of the original random variables:

$$\mathcal{F}_k = \sigma\left(\mathbf{gm}^{(1)}, \ldots, \mathbf{gm}^{(k)}\right).$$

Proof. We use induction on k. $k = 1$ is evident as $\mathbf{m}^{(1)} = \sqrt{q}\mathbf{1}$ and therefore $\mathbf{gm}^{(1)} = \sqrt{q}\xi^{(1)}$.

We assume that the equation is correct for $k \geq 1$. $\mathbf{g}^{(k+1)}\mathbf{M}^{(k+1)} = \left\|\mathbf{M}^{(k+1)}\right\|\mathbf{g}^{(k+1)}\phi^{(k+1)} = \left\|\mathbf{M}^{(k+1)}\right\|\xi^{(k+1)}$, and in the proof of (3.4), we have seen that $\mathbf{g}^{(k+1)}\mathbf{M}^{(k+1)} = \mathbf{gm}^{(k+1)} + \mathbf{R}^{(k)}$, where $\mathbf{R}^{(k)}$ is $\mathcal{F}_k$-measurable. Therefore

$$\mathbf{gm}^{(k+1)} + \mathbf{R}^{(k)} = \left\|\mathbf{M}^{(k+1)}\right\|\xi^{(k+1)}.$$

$\left\|\mathbf{M}^{(k+1)}\right\|$ is $\mathcal{F}_k$-measurable. Remember also that we had defined $\phi^{(k+1)} = 1$ on the $\mathcal{F}_k$-measurable event $\left\{\left\|\mathbf{M}^{(k+1)}\right\| = 0\right\}$. From that it follows that

$$\mathcal{F}_{k+1} = \sigma\left(\mathcal{F}_k, \xi^{(k+1)}\right) = \sigma\left(\mathcal{F}_k, \mathbf{gm}^{(k+1)}\right),$$

which by the induction hypothesis for k proves

$$\mathcal{F}_{k+1} = \sigma\left(\mathbf{gm}^{(1)}, \ldots, \mathbf{gm}^{(k+1)}\right).$$

□

Remark 3.5. The $\mathbf{g}^{(k)}$ can of course be expressed "explicitly" in terms of the $\mathbf{g}$, but already the expression for $\mathbf{g}^{(2)}$ is fairly involved. We haven't found a structurally simple formula for the $\mathbf{g}^{(k)}$ for general k. If it would exist, it would probably considerably simplify the arguments in the next section, but we doubt that there is one.

4. Computation of the Conditional Covariances of $\mathbf{g}^{(k)}$

We introduce some more notations.

We write $O_k\left(N^{-r}\right)$ for a generic sequence $\left\{X^{(N)}\right\}$ of $\mathcal{F}_k$-measurable non negative random variables which satisfies

$$\mathbb{P}\left(N^r X^{(N)} \geq K\right) \leq C \exp\left[-N/C\right],$$

for some $K, C > 0$. As usual, we don't write N as an index in the random variables. The constants C, $K > 0$ here may depend on h, β, and the level k, and on the formula where they appear, but on nothing else, in particular not on N, and any further indices. For instance, if we write

$$X_{ij} = Y_{ij} + O_k\left(N^{-5}\right),$$

we mean that there exists $C\left(\beta, h, k\right),\ K\left(\beta, h, k\right) > 0$ with

$$\sup_{ij} \mathbb{P}\left(N^5 \left|X_{ij} - Y_{ij}\right| \geq K\right) \leq C \exp\left[-N/C\right].$$

Furthermore, in such a case, it is tacitly assumed that $X_{ij} - Y_{ij}$ are $\mathcal{F}_k$-measurable.

Evidently, if X, Y are $O_k\left(N^{-r}\right)$, then $X + Y$ is $O_k\left(N^{-r}\right)$, and if X is $O_k\left(N^{-r}\right)$, and Y is $O_k\left(N^{-s}\right)$, then XY is $O_k\left(N^{-r-s}\right)$.

We write $\mathbb{E}_k$ for the conditional expectation, given $\mathcal{F}_k$.

We will finally prove the validity of the following relations:

$$\mathbb{E}_{k-2} g_{ij}^{(k)2} = \frac{1}{N} + O_{k-2}\left(N^{-2}\right), \tag{4.1}$$

$$\mathbb{E}_{k-2} g_{ij}^{(k)} g_{jt}^{(k)} = -\sum_{m=1}^{k-1} \frac{\phi_i^{(m)} \phi_t^{(m)}}{N^2} + O_{k-2}\left(N^{-3}\right), \quad \forall t \neq i, j, \tag{4.2}$$

$$\mathbb{E}_{k-2} g_{ij}^{(k)} g_{st}^{(k)} = \frac{\alpha_{ijst}^{(k)}}{N^3} + O_{k-2}\left(N^{-4}\right), \quad \text{if } \{s, t\} \cap \{i, j\} = \emptyset, \tag{4.3}$$

where

$$\alpha_{ijst}^{(k)} = \sum_{m=1}^{k-1} \sum_{A \subset \{i,j,s,t\}} \lambda_{m,A}^{(k)} \phi_A^{(m)}$$

with

$$\phi_A^{(m)} \stackrel{\text{def}}{=} \prod_{u \in A} \phi_u^{(m)}.$$

The $\lambda_{m,A}^{(k)}$ are real numbers, not random variables, which depend on A only through the type of subset which is taken with respect to the two subsets $\{i, j\}$, $\{s, t\}$. More precisely, there is only one number for $|A| = 4$, one for $|A| = 3$, and one for $|A| = 3$, but possibly two if $|A| = 2$, namely one for $A = \{i, j\}$ or $A = \{s, t\}$, and one for the other cases $\{i, s\}$, $\{i, t\}$, $\{j, s\}$, $\{j, t\}$. So totally, for any k, m, there are five possible λ's.

The main result of this section is:

Proposition 4.1. *Let $J \in \mathbb{N}$, assume* COND (J)*, and assume the validity of* (4.1)–(4.3) *hold for $k \leq J$. Then they hold for $k = J + 1$.*

The main point with assuming COND (J) is (2.11). On A_J, the variables $\phi^{(k)}$ are bounded for $k \leq J$.

Lemma 4.2. *Assume* (4.1)–(4.3) *for $k = J$, and* (2.11)*. Then*

(a)

$$\mathbb{E}_{J-1} \xi_i^{(J)2} = 1 + O_{J-1}\left(N^{-1}\right). \tag{4.4}$$

(b)

$$\mathbb{E}_{J-1} \xi_i^{(J)} \xi_j^{(J)} = \frac{1}{N} \phi_i^{(J)} \phi_j^{(J)} - \frac{1}{N} \sum_{r=1}^{J-1} \phi_i^{(r)} \phi_j^{(r)} + O_{J-1}\left(N^{-2}\right). \tag{4.5}$$

(c)

$$\mathbb{E}_{J-1} g_{ij}^{(J)} \xi_i^{(J)} = \frac{\phi_j^{(J)}}{N} + O_{J-1}\left(N^{-2}\right). \tag{4.6}$$

(d) *For $s \neq i, j$,*

$$\mathbb{E}_{J-1} g_{ij}^{(J)} \xi_s^{(J)} = -\frac{\phi_i^{(J)}}{N^2} \sum_{m=1}^{J-1} \phi_j^{(m)} \phi_s^{(m)} - \frac{\phi_j^{(J)}}{N^2} \sum_{m=1}^{J-1} \phi_i^{(m)} \phi_s^{(m)} + O_{J-1}\left(N^{-3}\right). \tag{4.7}$$

Proof. (a) As $\phi^{(J)}$ is $\mathcal{F}_{J-1}$-measurable, and $\mathbf{g}^{(J)}$ is independent of $\mathcal{F}_{J-1}$, conditionally on $\mathcal{F}_{J-2}$, we get

$$\begin{aligned}\mathbb{E}_{J-1}\xi_i^{(J)2} &= \sum_{s,t\neq i}\phi_s^{(J)}\phi_t^{(J)}\mathbb{E}_{J-1}\left(g_{is}^{(J)}g_{it}^{(J)}\right) = \sum_{s,t\neq i}\phi_s^{(J)}\phi_t^{(J)}\mathbb{E}_{J-2}\left(g_{is}^{(J)}g_{it}^{(J)}\right)\\ &= \sum_{s\neq i}\phi_s^{(J)2}\mathbb{E}_{J-2}\left(g_{is}^{(J)2}\right) + \sum_{\substack{s,t\neq i\\ s\neq t}}\phi_s^{(J)}\phi_t^{(J)}\mathbb{E}_{J-2}\left(g_{is}^{(J)}g_{it}^{(J)}\right).\end{aligned}$$

Using (4.1), (4.2), and the boundedness of the ϕ's on A_J, and $N^{-1}\sum_i \phi_i^{(J)2} = 1$, $\sum_i \phi_i^{(J)}\phi_i^{(m)} = 0$ for $m < J$, we get

$$\mathbb{E}_{J-1}\xi_i^{(J)2} = 1 + O_{J-1}\left(N^{-1}\right).$$

(b)

$$\mathbb{E}_{J-1}\xi_i^{(J)}\xi_j^{(J)} = \sum_{s\neq i,t\neq j}\phi_s^{(J)}\phi_t^{(J)}\mathbb{E}_{J-2}\left(g_{is}^{(J)}g_{jt}^{(J)}\right).$$

We split the sum over (s,t) into the one summand $s = j, t = i$, in $A = \{(s,s) : s \neq i, j\}$, $B = \{(j,t) : t \neq i, j\}$, $C = \{(s,i) : s \neq i, j\}$, and $D = \{(s,t) : \{s,t\} \cap \{i,j\} = \emptyset\}$. The one summand $s = j, t = i$ gives $\phi_i^{(J)}\phi_j^{(J)}/N + O_{J-1}\left(N^{-2}\right)$,

$$\begin{aligned}\sum_A &= \sum_{s\neq i,j}\phi_s^{(J)2}\mathbb{E}_{J-2}\left(g_{is}^{(J)}g_{js}^{(J)}\right) = \sum_{s\neq i,j}\phi_s^{(J)2}\left\{-\sum_{m=1}^{J-1}\frac{\phi_i^{(m)}\phi_j^{(m)}}{N^2} + O_{J-2}\left(N^{-3}\right)\right\}\\ &= -\sum_{m=1}^{J-1}\frac{\phi_i^{(m)}\phi_j^{(m)}}{N} + O_{J-1}\left(N^{-2}\right).\end{aligned}$$

$$\begin{aligned}\sum_B &= \sum_{t\neq i,j}\phi_j^{(J)}\phi_t^{(J)}\mathbb{E}_{J-2}\left(g_{ij}^{(J)}g_{jt}^{(J)}\right)\\ &= \sum_{t\neq i,j}\phi_j^{(J)}\phi_t^{(J)}\left\{-\sum_{m=1}^{J-1}\frac{\phi_i^{(m)}\phi_t^{(m)}}{N^2} + O_{J-2}\left(N^{-3}\right)\right\}.\end{aligned}$$

Because $\langle\phi^{(J)}, \phi^{(m)}\rangle = 0$ for $m < J$, this is seen to be $O_{J-1}\left(N^{-2}\right)$. The same applies to $\sum_C$.

It remains to consider the last part $\sum_D$. Here we have to use the expression for $\mathbb{E}_{J-2}\left(g_{ij}^{(J)}g_{st}^{(J)}\right)$ where $\{i,j\}\cap\{s,t\} = \emptyset$ given by (4.3),

$$\begin{aligned}&\sum_{s,t:\{s,t\}\cap\{i,j\}=\emptyset}\phi_s^{(J)}\phi_t^{(J)}\left[\frac{1}{N^3}\sum_{m=1}^{J-1}\sum_{A\subset\{i,j,s,t\}}\lambda_{m,A}^{(J)}\phi_A^{(m)} + O_{J-2}\left(N^{-4}\right)\right]\\ &= \frac{1}{N^3}\sum_{s,t:\{s,t\}\cap\{i,j\}=\emptyset}\phi_s^{(J)}\phi_t^{(J)}\sum_{m=1}^{J-1}\sum_{A\subset\{i,j,s,t\}}\lambda_{m,A}^{(J)}\phi_A^{(m)} + O_{J-2}\left(N^{-2}\right).\end{aligned}$$

Take e.g. $A = \{i, j, s\}$. Then $\lambda_{m,A}^{(J)} = \lambda_{m,3}^{(J)}$ with no further dependence of this number on i, j, s. So we get for this part for any summand on m with $m < J$,

$$\frac{1}{N^3}\lambda_{m,3}^{(J)} \sum_{s,t:\{s,t\}\cap\{i,j\}=\emptyset} \phi_s^{(J)}\phi_t^{(J)}\phi_s^{(m)}\phi_i^{(m)}\phi_j^{(m)}.$$

Using again $\left\langle \phi^{(J)}, \phi^{(m)} \right\rangle = 0$, we get that this is $O_{J-1}\left(N^{-2}\right)$. This applies in the same way to all the parts. Therefore (b) follows.

(c)

$$\begin{aligned}\mathbb{E}_{J-1} g_{ij}^{(J)} \xi_i^{(J)} &= \sum_{t \neq i} \phi_t^{(J)} \mathbb{E}_{J-2}\left(g_{ij}^{(J)} g_{it}^{(J)}\right) = \frac{\phi_j^{(J)}}{N} + O_{J-1}\left(N^{-2}\right) \\ &\quad + \sum_{t \neq i,j} \phi_t^{(J)} \left[-\sum_{m=1}^{J-1} \frac{\phi_j^{(m)}\phi_t^{(m)}}{N^2}\right] + O_{J-1}\left(N^{-2}\right) \\ &= \frac{\phi_j^{(J)}}{N} + O_{J-1}\left(N^{-2}\right),\end{aligned}$$

due to the orthogonality of the $\phi^{(m)}$.

(d)

$$\begin{aligned}\mathbb{E} g_{ij}^{(J)} \xi_s^{(J)} &= \sum_{t \neq s} \phi_t^{(J)} \mathbb{E} g_{ij}^{(J)} g_{st}^{(J)} \\ &= \phi_i^{(J)} \mathbb{E} g_{ij}^{(J)} g_{si}^{(J)} + \phi_j^{(J)} \mathbb{E} g_{ij}^{(J)} g_{sj}^{(J)} + O_{J-1}\left(N^{-3}\right),\end{aligned}$$

due again to (4.3). We therefore get

$$\mathbb{E} g_{ij}^{(J)} \xi_s^{(J)} = -\frac{1}{N^2} \sum_{m=1}^{J-1} \phi_s^{(m)} \left[\phi_i^{(J)}\phi_j^{(m)} + \phi_j^{(J)}\phi_i^{(m)}\right] + O_{J-1}\left(N^{-3}\right).$$

□

Lemma 4.3. *We assume the same as in Lemma 4.2. Put*

$$\hat{g}_{ij}^{(J)} \stackrel{\text{def}}{=} g_{ij}^{(J)} - \frac{\phi_i^{(J)}\xi_j^{(J)} + \phi_j^{(J)}\xi_i^{(J)}}{N} + \phi_i^{(J)}\phi_j^{(J)} \frac{1}{N^2} \sum_{r=1}^{N} \phi_r^{(J)}\xi_r^{(J)}.$$

Then

$$g_{ij}^{(J+1)} = \hat{g}_{ij}^{(J)} - \sum_s x_{ij,s}^{(J)} \xi_s^{(J)}, \tag{4.8}$$

where the $\mathcal{F}_{J-1}$-measurable coefficients $x_{ij,s}^{(J)}$ satisfy

$$\sum_s x_{ij,s}^{(J)} \phi_s^{(m)} = 0, \quad \forall i, j, \ \forall m < J, \tag{4.9}$$

with

$$x_{ij,s}^{(J)} = O_{J-1}\left(N^{-2}\right), \quad s \in \{i,j\},$$
$$x_{ij,s}^{(J)} = O_{J-1}\left(N^{-3}\right), \quad s \notin \{i,j\}.$$

Proof. The existence of $\mathcal{F}_{J-1}$-measurable coefficients $x_{ij,s}^{(J)}$ comes from linear algebra.

Remark that

$$\sum_s \xi_s^{(J)} \phi_s^{(m)} = \sum_{s,j} \phi_s^{(m)} g_{sj}^{(J)} \phi_j^{(J)} = \sum_j \phi_j^{(J)} \left[\sum_s g_{js}^{(J)} \phi_s^{(m)}\right] = 0.$$

Therefore, we can replace the $x_{ij,\cdot}^{(J)}$ by

$$x_{ij,\cdot}^{(J)} - \sum_{m=1}^{J-1} \left\langle x_{ij,\cdot}^{(J)}, \phi^{(m)} \right\rangle \phi^{(m)},$$

which satisfy the desired property (4.9).

We keep i, j fixed for the moment and write x_s for $x_{ij,s}^{(J)}$. The requirement for them is that for all t,

$$\mathbb{E}_{J-1}\left(\left(\hat{g}_{ij}^{(J)} - \sum_s x_s \xi_s^{(J)}\right) \xi_t^{(J)}\right) = 0 \tag{4.10}$$

(see the definition of the $\mathbf{g}^{(J)}$ in (3.2)).

From Lemma 4.2, we get

$$\mathbb{E}_{J-1}\left(\hat{g}_{ij}^{(J)} \xi_i^{(J)}\right) = O_{J-1}\left(N^{-2}\right),$$

and the same for $\mathbb{E}_{J-1}\left(\hat{g}_{ij}^{(J)} \xi_j^{(J)}\right)$. For $t \notin \{i,j\}$, we have

$$\begin{aligned}
\mathbb{E}_{J-1}\left(\hat{g}_{ij}^{(J)} \xi_t^{(J)}\right) &= -\frac{\phi_i^{(J)}}{N^2} \sum_{m=1}^{J-1} \phi_j^{(m)} \phi_t^{(m)} - \frac{\phi_j^{(J)}}{N^2} \sum_{m=1}^{J-1} \phi_i^{(m)} \phi_t^{(m)} + O_{J-1}\left(N^{-3}\right) \\
&\quad - \frac{\phi_i^{(J)}}{N}\left\{\frac{1}{N}\phi_j^{(J)}\phi_t^{(J)} - \frac{1}{N}\sum_{m=1}^{J-1}\phi_j^{(m)}\phi_t^{(m)} + O_{J-1}\left(N^{-2}\right)\right\} \\
&\quad - \frac{\phi_j^{(J)}}{N}\left\{\frac{1}{N}\phi_i^{(J)}\phi_t^{(J)} - \frac{1}{N}\sum_{m=1}^{J-1}\phi_i^{(m)}\phi_t^{(m)} + O_{J-1}\left(N^{-2}\right)\right\} \\
&\quad + \phi_i^{(J)}\phi_j^{(J)}\frac{1}{N^2}\sum_r \phi_r^{(J)} \mathbb{E}_{J-1}\xi_r^{(J)}\xi_t^{(J)} \\
&= -\frac{2}{N^2}\phi_i^{(J)}\phi_j^{(J)}\phi_t^{(J)} + \frac{1}{N^2}\phi_i^{(J)}\phi_j^{(J)}\phi_t^{(J)} \\
&\quad + \frac{1}{N^2}\phi_i^{(J)}\phi_j^{(J)}\sum_{r \neq t}\phi_r^{(J)}\left\{\frac{1}{N}\phi_r^{(J)}\phi_t^{(J)} - \frac{1}{N}\sum_{m=1}^{J-1}\phi_r^{(m)}\phi_t^{(m)}\right\} \\
&\quad + O_{J-1}\left(N^{-3}\right).
\end{aligned}$$

Due to the orthonormality of the ϕ, one gets

$$\frac{1}{N}\sum_{r\neq t}\phi_r^{(J)2} = 1 + O_{J-1}\left(N^{-1}\right),$$
$$\sum_{r\neq t}\phi_r^{(J)}\phi_r^{(m)} = O_{J-1}\left(N^{-1}\right).$$

So we get

$$\mathbb{E}_{J-1}\left(\hat{g}_{ij}^{(J)}\xi_t^{(J)}\right) = O_{J-1}\left(N^{-3}\right).$$

We write for the moment $y_t \overset{\text{def}}{=} \mathbb{E}_{J-1}\left(\hat{g}_{ij}^{(J)}\xi_t^{(J)}\right)$. The equations (4.10) for $\{x_s\}$ are

$$\sum_s x_s \mathbb{E}_{J-1}\xi_s^{(J)}\xi_t^{(J)} = y_t,\ \forall t.$$

Writing r_{ij} for the $O_{J-1}\left(N^{-2}\right)$ error term in (4.5), and for $j = i$, the $O_{J-1}\left(N^{-1}\right)$ error term in (4.4), we arrive at

$$\sum_{s\neq t} x_s\left\{\frac{1}{N}\phi_s^{(J)}\phi_t^{(J)} - \frac{1}{N}\sum_{m=1}^{J-1}\phi_s^{(m)}\phi_t^{(m)} + r_{st}\right\} + x_t\left(1 + r_{tt}\right) = y_t.$$

In the first summand, we sum now over all s, remarking that we have assumed that $\sum_s x_s\phi_s^{(m)} = 0$ for $m < J$. The error for not summing over the single t can be incorporated into r_{tt}. We therefore arrive at

$$x_t + \phi_t^{(J)}\frac{1}{N}\sum_s x_s\phi_s^{(J)} + \sum_s x_s r_{st} = y_t.$$

Write Φ for the matrix $\left(N^{-1}\phi_i^{(J)}\phi_j^{(J)}\right)$ and R for $\left(r_{ij}\right)$. Then we have to invert the matrix $(I + \Phi + R)$. Remark that $(I + \Phi)^{-1} = I - \Phi/2$. Therefore

$$(I - \Phi/2)\,(I + \Phi + R) = I + (I - \Phi/2)\,R.$$

We can develop the right hand side as a Neumann series:

$$\begin{aligned}(I + \Phi + R)^{-1}(I + \Phi) &= (I + (I - \Phi/2)\,R)^{-1}\\ &= I - (I - \Phi/2)\,R + [(I - \Phi/2)\,R]^2 - \cdots\\ (I + \Phi + R)^{-1} &= I - \frac{\Phi}{2} - \left(I - \frac{\Phi}{2}\right)R\left(I - \frac{\Phi}{2}\right) + \cdots.\end{aligned}$$

As $(\Phi\mathbf{y})_i = O_{J-1}\left(N^{-3}\right)$, we get the desired conclusion. □

Proof of Proposition 4.1.

$$\begin{aligned}\mathbb{E}_{J-1}\left(g_{ij}^{(J+1)}g_{st}^{(J+1)}\right) &= \mathbb{E}_{J-1}\left(\hat{g}_{ij}^{(J+1)}\hat{g}_{st}^{(J+1)}\right) - \sum_u x_{st,u}^{(J)}\mathbb{E}_{J-1}\left(\xi_u^{(J)}\hat{g}_{ij}^{(J)}\right) \\ &\quad - \sum_u x_{ij,u}^{(J)}\mathbb{E}_{J-1}\left(\xi_u^{(J)}\hat{g}_{st}^{(J)}\right) \\ &\quad + \sum_{u,v} x_{ij,u}^{(J)}x_{st,v}^{(J)}\mathbb{E}_{J-1}\left(\xi_u^{(J)}\xi_v^{(J)}\right). \end{aligned} \tag{4.11}$$

The summands involving the $x^{(J)}$ all only give contributions which enter the O_{J-1}-terms. Take for instance $s = j,\ t \neq i, j$. In that case, the claimed O_{J-1}-term is $O_{J-1}\left(N^{-3}\right)$. In the last summand of (4.11), there is one summand, namely $u = v = j$, where the $x^{(J)}$ are $O_{J-1}\left(N^{-2}\right)$, so this summand is only $O_{J-1}\left(N^{-4}\right)$,

$$\begin{aligned}\sum_u x_{jt,u}^{(J)}\mathbb{E}_{J-1}\left(\xi_u^{(J)}\hat{g}_{ij}^{(J)}\right) &= x_{jt,i}^{(J)}\mathbb{E}_{J-1}\left(\xi_i^{(J)}\hat{g}_{ij}^{(J)}\right) + x_{jt,j}^{(J)}\mathbb{E}_{J-1}\left(\xi_j^{(J)}\hat{g}_{ij}^{(J)}\right) \\ &\quad + x_{jt,t}^{(J)}\mathbb{E}_{J-1}\left(\xi_t^{(J)}\hat{g}_{ij}^{(J)}\right) \\ &\quad + \sum_{u\neq i,j,t} x_{jt,u}^{(J)}\mathbb{E}_{J-1}\left(\xi_u^{(J)}\hat{g}_{ij}^{(J)}\right). \end{aligned} \tag{4.12}$$

From Lemma 4.2, we get

$$\begin{aligned}\mathbb{E}_{J-1}\left(\xi_i^{(J)}\hat{g}_{ij}^{(J)}\right) &= \mathbb{E}_{J-1}\left(\xi_i^{(J)}g_{ij}^{(J)}\right) - N^{-1}\phi_i^{(J)}\mathbb{E}_{J-1}\left(\xi_i^{(J)}\xi_j^{(J)}\right) \\ &\quad - N^{-1}\phi_j^{(J)}\mathbb{E}_{J-1}\left(\xi_i^{(J)2}\right) \\ &\quad + \phi_i^{(J)}\phi_j^{(J)}N^{-2}\sum_{r=1}^N \phi_r^{(J)}\mathbb{E}_{J-1}\left(\xi_i^{(J)}\xi_r^{(J)}\right) \\ &= O_{J-1}\left(N^{-1}\right) \end{aligned}$$

and similarly $\mathbb{E}_{J-1}\left(\xi_j^{(J)}\hat{g}_{ij}^{(J)}\right) = O_{J-1}\left(N^{-1}\right)$, and $\mathbb{E}_{J-1}\left(\xi_u^{(J)}\hat{g}_{ij}^{(J)}\right)$ for $u \notin \{i, j\}$. So the sum in (4.12) is

$$\begin{aligned}&O_{J-1}\left(N^{-3}\right)O_{J-1}\left(N^{-1}\right) + O_{J-1}\left(N^{-2}\right)O_{J-1}\left(N^{-1}\right) \\ &\quad + O_{J-1}\left(N^{-2}\right)O_{J-1}\left(N^{-2}\right) + N\,O_{J-1}\left(N^{-3}\right)O_{J-1}\left(N^{-2}\right) \\ &= O_{J-1}\left(N^{-3}\right). \end{aligned}$$

The other summands behave similarly. The third and fourth summand in (4.11) behave similarly.

As another case, take $\{i, j\} \cap \{s, t\} = \emptyset$, where we have to get $O_{J-1}\left(N^{-4}\right)$ for the second to fourth summand in (4.11),

$$\begin{aligned}
\sum_u x_{st,u}^{(J)} \mathbb{E}_{J-1}\left(\xi_u^{(J)} \hat{g}_{ij}^{(J)}\right) &= \sum_{u=i,j} + \sum_{u=s,t} + \sum_{u \notin \{i,j,s,t\}} \\
&= O_{J-1}\left(N^{-3}\right) O_{J-1}\left(N^{-1}\right) \\
&\quad + O_{J-1}\left(N^{-2}\right) O_{J-1}\left(N^{-2}\right) \\
&\quad + N O_{J-1}\left(N^{-3}\right) O_{J-1}\left(N^{-2}\right) \\
&= O_{J-1}\left(N^{-4}\right). \\
\sum_{u,v} x_{ij,u}^{(J)} x_{st,v}^{(J)} \mathbb{E}_{J-1}\left(\xi_u^{(J)} \xi_v^{(J)}\right) &= \sum_{u=v \in \{i,j,s,t\}} + \sum_{u=v \notin \{i,j,s,t\}} + \sum_{u \in \{i,j\}} \sum_{v \in \{s,t\}} \\
&\quad + \sum_{u \in \{i,j\}} \sum_{v \notin \{s,t\}, \neq u} + \sum_{v \in \{s,t\}} \sum_{u \notin \{i,j\}, \neq v} \\
&\quad + \sum_{u \neq v} \sum_{u \notin \{i,j\}} \sum_{v \notin \{s,t\}} \\
&= O_{J-1}\left(N^{-5}\right) + N O_{J-1}\left(N^{-6}\right) \\
&\quad + O_{J-1}\left(N^{-2}\right) O_{J-1}\left(N^{-2}\right) O_{J-1}\left(N^{-1}\right) \\
&\quad + N O_{J-1}\left(N^{-2}\right) O_{J-1}\left(N^{-3}\right) O_{J-1}\left(N^{-1}\right) \\
&\quad + N O_{J-1}\left(N^{-2}\right) O_{J-1}\left(N^{-3}\right) O_{J-1}\left(N^{-1}\right) \\
&\quad + N^2 O_{J-1}\left(N^{-3}\right) O_{J-1}\left(N^{-3}\right) O_{J-1}\left(N^{-1}\right) \\
&= O_{J-1}\left(N^{-5}\right),
\end{aligned}$$

which is better than required.

In order to prove (4.1)–(4.3), it therefore remains to investigate $\mathbb{E}_{J-1}\left(\hat{g}_{ij}^{(J+1)} \hat{g}_{st}^{(J+1)}\right)$. For (4.1):

$$\mathbb{E}_{J-1}\left(\hat{g}_{ij}^{(J)2}\right) = \mathbb{E}_{J-1}\left[\left(g_{ij}^{(J)} - \frac{\phi_i^{(J)} \xi_j^{(J)} + \phi_j^{(J)} \xi_i^{(J)}}{N} + \frac{\phi_i^{(J)} \phi_j^{(J)}}{N^2} \sum_t \phi_t^{(J)} \xi_t^{(J)}\right)^2\right].$$

Using Lemma 4.2, one easily gets that anything except $\mathbb{E}_{J-1}\left(g_{ij}^{(J)2}\right)$ is $O_{J-1}\left(N^{-2}\right)$. $\mathbb{E}_{J-1}\left(g_{ij}^{(J)2}\right) = \mathbb{E}_{J-2}\left(g_{ij}^{(J)2}\right)$ from the conditional independence of $\mathbf{g}^{(J)}$ of $\mathcal{F}_{J-1}$, given $\mathcal{F}_{J-2}$. So the claim follows.

For (4.2):

$$\begin{aligned}
\mathbb{E}_{J-1}\left(\hat{g}_{ij}^{(J)} \hat{g}_{jt}^{(J)}\right) = \mathbb{E}_{J-1}\Bigg[&\left(g_{ij}^{(J)} - \frac{\phi_i^{(J)} \xi_j^{(J)} + \phi_j^{(J)} \xi_i^{(J)}}{N} + \frac{\phi_i^{(J)} \phi_j^{(J)}}{N^2} \sum_u \phi_u^{(J)} \xi_u^{(J)}\right) \\
&\times \left(g_{jt}^{(J)} - \frac{\phi_j^{(J)} \xi_t^{(J)} + \phi_t^{(J)} \xi_j^{(J)}}{N} + \frac{\phi_j^{(J)} \phi_t^{(J)}}{N^2} \sum_u \phi_u^{(J)} \xi_u^{(J)}\right)\Bigg].
\end{aligned}$$

We write $m \times n$ for the summand, we get by multiplying the m^{th} summand in the first bracket with the n^{th} in the second. By induction hypothesis, we get

$$1 \times 1 = -\sum_{m=1}^{J-1} \frac{\phi_i^{(m)}\phi_t^{(m)}}{N^2} + O_{J-2}\left(N^{-3}\right).$$

In the 1×2-term, only the multiplication of $g_{ij}^{(J)}$ with $\xi_j^{(J)}$ counts, the other part giving $O_{J-2}\left(N^{-3}\right)$. Therefore

$$\begin{aligned} 1 \times 2 &= -\frac{\phi_t^{(J)}}{N}\mathbb{E}_{J-1} g_{ij}^{(J)}\xi_j^{(J)} + O_{J-1}\left(N^{-3}\right) \\ &= -\frac{\phi_i^{(J)}\phi_t^{(J)}}{N^2} + O_{J-1}\left(N^{-3}\right). \end{aligned}$$

2×1 gives the same. In 2×2, again only the matching of $\xi_j^{(J)}$ with $\xi_j^{(J)}$ counts, so we get

$$2 \times 2 = \frac{\phi_i^{(J)}\phi_t^{(J)}}{N^2} + O_{J-1}\left(N^{-3}\right).$$

The other parts are easily seen to give $O_{J-1}\left(N^{-3}\right)$. We have proved that

$$\mathbb{E}_{J-1}\left(\hat{g}_{ij}^{(J)}\hat{g}_{jt}^{(J)}\right) = -\sum_{m=1}^{J} \frac{\phi_i^{(m)}\phi_t^{(m)}}{N^2} + O_{J-1}\left(N^{-3}\right).$$

Finally for (4.3), we have here $\{i, j\} \cap \{s, t\} = \emptyset$,

$$\begin{aligned} \mathbb{E}_{J-1}\left(\hat{g}_{ij}^{(J)}\hat{g}_{st}^{(J)}\right) = \mathbb{E}_{J-1}&\left(g_{ij}^{(J)} - \frac{\phi_i^{(J)}\xi_j^{(J)} + \phi_j^{(J)}\xi_i^{(J)}}{N} + \frac{\phi_i^{(J)}\phi_j^{(J)}}{N^2}\sum_u \phi_u^{(J)}\xi_u^{(J)}\right) \\ &\times\left(g_{st}^{(J)} - \frac{\phi_s^{(J)}\xi_t^{(J)} + \phi_t^{(J)}\xi_s^{(J)}}{N} + \frac{\phi_s^{(J)}\phi_t^{(J)}}{N^2}\sum_u \phi_u^{(J)}\xi_u^{(J)}\right). \end{aligned}$$

The 1×1, 1×2, 2×1, and 2×2-terms are clearly of the desired form, either from induction hypothesis or Lemma 4.2,

$$1 \times 3 = \frac{\phi_s^{(J)}\phi_t^{(J)}}{N^2}\sum_u \mathbb{E}_{J-1}\left(g_{ij}^{(J)}\phi_u^{(J)}\xi_u^{(J)}\right).$$

For $u = i$ we get for the expectation $\phi_i^{(J)}\phi_j^{(J)}/N + O_{J-1}\left(N^{-2}\right)$, so this is of the desired form. The same applies to $u = j$. It therefore remains

$$\begin{aligned} &\frac{\phi_s^{(J)}\phi_t^{(J)}}{N^2}\sum_{u\neq i,j} \phi_u^{(J)}\mathbb{E}_{J-1}\left(g_{ij}^{(J)}\xi_u^{(J)}\right) \\ &= \frac{\phi_s^{(J)}\phi_t^{(J)}}{N^2}\sum_{u\neq i,j} \phi_u^{(J)}\left\{-\frac{1}{N^2}\sum_{m=1}^{J-1}\phi_u^{(m)}\left[\phi_i^{(J)}\phi_j^{(m)} + \phi_j^{(J)}\phi_i^{(m)}\right]\right\} + O_{J-1}\left(N^{-4}\right). \end{aligned}$$

As $\sum_u \phi_u^{(J)}\phi_u^{(m)} = 0$, the whole expression is $O_{J-1}\left(N^{-4}\right)$. The other cases are handled similarly. □

5. Proof of Proposition 2.5

Proof of the proposition. The proof is short given the results of two main lemmas which are formulated and proved afterwards.

We assume COND (J), and (4.1)–(4.3) for $k \leq J$. By Proposition 4.1 of the last section, this implies (4.1)–(4.3) for $k \leq J+1$. Using this, we prove now (2.5) and (2.6) for $k = J+1$, so that we have proved COND $(J+1)$. Having achieved this, the proof of Proposition 2.5 is complete.

We have to introduce some more notations:

For $j < k$, define $\mathbf{X}^{(k,0)} \stackrel{\text{def}}{=} 0$, $\mathbf{X}^{(k,j)} \stackrel{\text{def}}{=} \sum_{t=1}^{j} \gamma_t \phi^{(t)}$, and $\mathbf{X}^{(k,k)} \stackrel{\text{def}}{=} \sum_{t=1}^{k-1} \gamma_t \phi^{(t)} + \sqrt{q - \Gamma_{k-1}^2}\phi^{(k)}$.

Remark that under COND (J),

$$\mathbf{m}^{(k)} \approx \mathbf{X}^{(k,k)}, \tag{5.1}$$

for $k \leq J$. Indeed

$$\begin{aligned} &\left\| \mathbf{m}^{(k)} - \sum_{m=1}^{k-1} \left\langle \mathbf{m}^{(k)}, \phi^{(m)} \right\rangle \phi^{(m)} \right\| \phi^{(k)} \\ &= \mathbf{m}^{(k)} - \sum_{m=1}^{k-1} \left\langle \mathbf{m}^{(k)}, \phi^{(m)} \right\rangle \phi^{(m)} \\ &\approx \mathbf{m}^{(k)} - \sum_{m=1}^{k-1} \gamma_m \phi^{(m)}. \end{aligned}$$

From $q > \Gamma_{k-1}^2$, by (2.5) and (2.6) for $k \leq J$, and the fact that the $\phi_j^{(k)}$ are uniformly bounded on A_J, we have

$$\left\| \mathbf{m}^{(k)} - \sum_{m=1}^{k-1} \left\langle \mathbf{m}^{(k)}, \phi^{(m)} \right\rangle \phi^{(m)} \right\| \simeq \sqrt{q - \Gamma_{k-1}^2}.$$

So the claim (5.1) follows.

We set for $1 \leq s < k$,

$$\mathbf{m}^{(k,s)} \stackrel{\text{def}}{=} \mathrm{Th}\left(\mathbf{g}^{(s)} \mathbf{M}^{(k-1,s-1)} + \sum_{t=1}^{s-1} \gamma_t \xi^{(t)} + \beta(1-q) \left\{ \mathbf{X}^{(k-2,s-1)} - \mathbf{m}^{(k-2)} \right\} \right).$$

Remark that by Lemma 3.2, we have $\mathbf{g}^{(s)} \mathbf{M}^{(k-1,s-1)} = \mathbf{g}^{(s)} \mathbf{m}^{(k-1)}$. Evidently

$$\mathbf{m}^{(k,1)} = \mathbf{m}^{(k)},$$

and we define

$$\hat{\mathbf{m}}^{(1)} \stackrel{\text{def}}{=} \mathbf{m}^{(1)} = \sqrt{q}\mathbf{1}, \quad \hat{\mathbf{m}}^{(k)} \stackrel{\text{def}}{=} \mathbf{m}^{(k,k-1)}, \quad k \geq 2.$$

By Lemma 5.1 below COND (J) implies

$$\mathbf{m}^{(J+1)} \approx \hat{\mathbf{m}}^{(J+1)}. \tag{5.2}$$

As COND (J) implies trivially COND $\left(J'\right)$ for $J' < J$, it follows that COND (J) implies $\mathbf{m}^{(k)} \approx \hat{\mathbf{m}}^{(k)}$ for all $k \leq J+1$. As the $m_j^{(k)}$ are uniformly bounded by 1, we get from that

$$\left\langle \mathbf{m}^{(J+1)}, \mathbf{m}^{(j)} \right\rangle \simeq \left\langle \hat{\mathbf{m}}^{(J+1)}, \hat{\mathbf{m}}^{(j)} \right\rangle,$$

for all $j \leq J+1$.

By Lemma 5.3 below, we also have

$$\left\langle \hat{\mathbf{m}}^{(J+1)}, \hat{\mathbf{m}}^{(j)} \right\rangle \simeq \rho_j$$

for $j \leq J$, and

$$\left\| \hat{\mathbf{m}}^{(J+1)} \right\|^2 \simeq q.$$

This implies COND $(J+1)$ which finishes the proof of Proposition 2.5. □

Unfortunately, there is a slight complication when proving (5.2), namely that we have to extend the induction scheme by proving in parallel

$$\left\langle \xi^{(m)}, \mathbf{m}^{(k)} \right\rangle \simeq \left\langle \xi^{(m)}, \hat{\mathbf{m}}^{(k)} \right\rangle, \quad \forall m < k \tag{5.3}$$

for $k = J+1$ which is not evident from (5.2) as the $\xi_i^{(m)}$ are not bounded.

Lemma 5.1. *Assume the validity of* (2.5)–(2.7) *and* (5.3) *for* $k \leq J$. *Then for* $s = 1, \ldots, J-1$,

$$\mathbf{m}^{(J+1,s)} \approx \mathbf{m}^{(J+1,s+1)},$$

and (5.3) holds for $k = J+1$. *In particular,*

$$\mathbf{m}^{(J+1)} \approx \hat{\mathbf{m}}^{(J+1)} \textit{ follows}.$$

Proof. We prove by induction on s, $1 \leq s \leq J-1$, that

$$\mathbf{m}^{(J+1,s)} \approx \mathbf{m}^{(J+1,s+1)}, \tag{5.4}$$

and

$$\left\langle \xi^{(m)}, \mathbf{m}^{(J+1,s)} \right\rangle \simeq \left\langle \xi^{(m)}, \mathbf{m}^{(J+1,s+1)} \right\rangle, \quad m \leq J. \tag{5.5}$$

We have

$$\mathbf{g}^{(s+1)} = \mathbf{g}^{(s)} - \xi^{(s)} \otimes_{\mathrm{s}} \phi^{(s)} + \left\langle \xi^{(s)}, \phi^{(s)} \right\rangle \left(\phi^{(s)} \otimes \phi^{(s)} \right) + \mathbf{c}^{(s)},$$

where

$$c_{ij}^{(s)} = \sum_r x_{ij,r}^{(s)} \xi_r^{(s)},$$

see Lemma 4.3. Therefore

$$\mathbf{m}^{(J+1,s)} = \mathrm{Th}\Big(\mathbf{g}^{(s+1)}\mathbf{m}^{(J)} + \mathbf{y} + \sum_{t=1}^{s-1}\gamma_t\xi^{(t)} + \beta(1-q)\Big\{\mathbf{X}^{(J-1,s-1)} - \mathbf{m}^{(J-1)}\Big\}\Big), \tag{5.6}$$

where

$$\mathbf{y} \stackrel{\text{def}}{=} \Big\langle\phi^{(s)}, \mathbf{m}^{(J)}\Big\rangle\xi^{(s)} + \Big\langle\xi^{(s)}, \mathbf{m}^{(J)}\Big\rangle\phi^{(s)} + \Big\langle\phi^{(s)}, \mathbf{m}^{(J)}\Big\rangle\Big\langle\phi^{(s)}, \xi^{(J)}\Big\rangle\phi^{(s)} + \mathbf{c}^{(s)}\mathbf{m}^{(J)}.$$

We write

$$\mathbf{y}^{(1)} \stackrel{\text{def}}{=} \Big\langle\phi^{(s)}, \mathbf{m}^{(J)}\Big\rangle\xi^{(s)} + \Big\langle\xi^{(s)}, \mathbf{m}^{(J)}\Big\rangle\phi^{(s)} + \Big\langle\phi^{(s)}, \mathbf{m}^{(J)}\Big\rangle\Big\langle\phi^{(s)}, \xi^{(J)}\Big\rangle\phi^{(s)},$$
$$\mathbf{y}^{(2)} \stackrel{\text{def}}{=} \Big\langle\phi^{(s)}, \mathbf{m}^{(J)}\Big\rangle\xi^{(s)} + \Big\langle\xi^{(s)}, \mathbf{m}^{(J)}\Big\rangle\phi^{(s)},$$
$$\mathbf{y}^{(3)} \stackrel{\text{def}}{=} \gamma_s\xi^{(s)} + \Big\langle\xi^{(s)}, \mathbf{m}^{(J)}\Big\rangle\phi^{(s)},$$
$$\mathbf{y}^{(4)} \stackrel{\text{def}}{=} \gamma_s\xi^{(s)} + \Big\langle\xi^{(s)}, \hat{\mathbf{m}}^{(J)}\Big\rangle\phi^{(s)},$$
$$\mathbf{y}^{(5)} \stackrel{\text{def}}{=} \begin{cases} \gamma_s\xi^{(s)} + \beta(1-q)\gamma_s\phi^{(s)} & \text{if } s < J-1 \\ \gamma_s\xi^{(s)} + \beta(1-q)\sqrt{q - \Gamma_{s-1}^2}\,\phi^{(s)} & \text{if } s = J-1 \end{cases},$$

and then set ad hoc

$$\mu^{(0)} \stackrel{\text{def}}{=} \mathbf{m}^{(J+1,s)},$$

and define $\mu^{(n)}$ where $\mathbf{y}$ is replaced by $\mathbf{y}^{(n)}$, $n = 1, \ldots, 5$. Remark that from (5.6),

$$\mu^{(5)} = \mathbf{m}^{(J+1,s+1)}.$$

We will prove

$$\mu^{(n-1)} \approx \mu^{(n)}, \quad n = 1, \ldots, 5, \tag{5.7}$$

and

$$\Big\langle\xi^{(m)}, \mu^{(n-1)}\Big\rangle \simeq \Big\langle\xi^{(m)}, \mu^{(n)}\Big\rangle, \quad n = 1, \ldots, 5. \tag{5.8}$$

which prove the desired induction in s.

To switch from $\mu^{(0)}$ to $\mu^{(1)}$, we observe that by the estimates of Lemma 4.3, one has

$$\Big|\Big(\mathbf{c}^{(s)}\mathbf{m}^{(J)}\Big)_i\Big| \leq O_{s-1}(1)\left[\frac{1}{N}\Big|\xi_i^{(s)}\Big| + \frac{1}{N^2}\sum_j\Big|\xi_j^{(s)}\Big|\right].$$

Therefore

$$\frac{1}{N}\sum_i\Big|\mu_i^{(0)} - \mu_i^{(1)}\Big| \leq \frac{O_{s-1}(1)}{N^2}\sum_j\Big|\xi_j^{(s)}\Big|,$$

and

$$\frac{1}{N}\sum_i \left|\xi_i^{(m)}\left(\mu_i^{(0)}-\mu_i^{(1)}\right)\right|$$
$$\leq \frac{O_{s-1}(1)}{N}\left\{\frac{1}{N}\sum_i\left|\xi_i^{(m)}\xi_i^{(s)}\right| + \frac{1}{N}\sum_i\left|\xi_i^{(m)}\right|\frac{1}{N}\sum_i\left|\xi_i^{(s)}\right|\right\}.$$

By choosing K large enough, we get for $1/\sqrt{N} \leq t \leq 1$ by Corollary A.2 (a),

$$\mathbb{P}\left(\frac{1}{N}\sum_i\left|\mu_i^{(0)}-\mu_i^{(1)}\right| \geq t\right) \leq \mathbb{P}\left(\frac{K}{N}\sum_j\left|\xi_j^{(s)}\right| \geq tN\right) + \mathbb{P}\left(O_{s-1}(1) \geq K\right)$$
$$\leq C\exp\left[-N/C\right] \leq C\exp\left[-Nt^2/C\right].$$

For $t \leq 1/\sqrt{N}$, the bound is trivial anyway. This proves (5.7) for $n=1$. Equation (5.8) follows in the same way using Corollary A.2 (b),

$$\frac{1}{N}\sum_i\left|\mu_i^{(1)}-\mu_i^{(2)}\right| \leq C\left|\left\langle\phi^{(s)},\mathbf{m}^{(J)}\right\rangle\left\langle\phi^{(s)},\xi^{(J)}\right\rangle\left\langle\phi^{(s)},\mathbf{1}\right\rangle\right|$$
$$\leq C\left|\left\langle\phi^{(s)},\xi^{(J)}\right\rangle\right|$$

on A_J. Equation (5.7) for $n=2$ then follows from Corollary A.2 (c). As for (5.8), we remark that

$$\frac{1}{N}\sum_i\left|\xi_i^{(m)}\left(\mu_i^{(1)}-\mu_i^{(2)}\right)\right| \leq C\left|\left\langle\phi^{(s)},\xi^{(J)}\right\rangle\right|\left|\left\langle\phi^{(s)},\xi^{(m)}\right\rangle\right|.$$

We can then again use Corollary A.2 (c) remarking that $\exp\left[-Nt/C\right] \leq \exp\left[-Nt^2/C\right]$ for $t \leq 1$,

$$\frac{1}{N}\sum_i\left|\mu_i^{(2)}-\mu_i^{(3)}\right| \leq C\left|\left\langle\phi^{(s)},\mathbf{m}^{(J)}\right\rangle-\gamma_s\right|\frac{1}{N}\sum_i\left|\xi_i^{(s)}\right|.$$

Equation (5.7) for $n=3$ follows from the induction hypothesis (2.7), and Corollary A.2 (a). Similarly with (5.8) but here, one has to use part (b) of Corollary A.2,

$$\frac{1}{N}\sum_i\left|\mu_i^{(3)}-\mu_i^{(4)}\right| \leq C\left|\left\langle\xi^{(s)},\mathbf{m}^{(J)}-\hat{\mathbf{m}}^{(J)}\right\rangle\right|$$

on A_k, and one uses the induction hypothesis (5.3) for J to get (5.7) for $n=4$. Remark that actually, one has a bound uniform in i :

$$\left|\mu_i^{(3)}-\mu_i^{(4)}\right| \leq C\left|\left\langle\xi^{(s)},\mathbf{m}^{(J)}-\hat{\mathbf{m}}^{(J)}\right\rangle\right|.$$

Therefore, one also gets (5.8) using Corollary A.2. Up to now, we have obtained

$$\mathbf{m}^{(J+1,s)} \approx \mathrm{Th}\left(\mathbf{g}^{(s)}\mathbf{M}^{(k-1,s-1)}+\sum_{t=1}^{s}\gamma_t\xi^{(t)}\right.$$
$$\left.+\left\langle\xi^{(s)},\hat{\mathbf{m}}^{(J)}\right\rangle\phi^{(s)}+\beta(1-q)\left\{\mathbf{X}^{(J-1,s-1)}-\mathbf{m}^{(J-1)}\right\}\right)$$

and

$$\left\langle \xi^{(m)}, \mathbf{m}^{(J+1,s)} \right\rangle \simeq \left\langle \xi^{(m)}, \mathrm{Th}\left(\mathbf{g}^{(s)} \mathbf{M}^{(k-1,s-1)} + \sum_{t=1}^{s} \gamma_t \xi^{(t)} \right.\right.$$
$$\left.\left. + \left\langle \xi^{(s)}, \hat{\mathbf{m}}^{(J)} \right\rangle \phi^{(s)} + \beta (1-q) \left\{ \mathbf{X}^{(J-1,s-1)} - \mathbf{m}^{(J-1)} \right\} \right) \right\rangle.$$

By Lemma 5.3 (a) below, we have

$$\left\langle \xi^{(s)}, \hat{\mathbf{m}}^{(J)} \right\rangle \simeq \begin{cases} \beta (1-q) \gamma_s & \text{for } s < J-1 \\ \beta (1-q) \sqrt{q - \Gamma_{J-2}^2} & \text{for } s = J-1 \end{cases}, \tag{5.9}$$

and we can therefore replace $\langle \xi^{(s)}, \hat{\mathbf{m}}^{(J)} \rangle \phi^{(s)}$ on the right hand side, by $\beta (1-q) \gamma_s \phi^{(s)}$ for $s < J-1$, or $\beta (1-q) \sqrt{q - \Gamma_{J-2}^2} \phi^{(J-1)}$ for $s = J-1$, which is the same as replacing $\mathbf{X}^{(J-1,s-1)}$ by $\mathbf{X}^{(J-1,s)}$. Therefore, the lemma is proved. □

Remark 5.2. $\hat{\mathbf{m}}^{(k)}$ still contains inside $\mathrm{Th}(\cdot)$ the summand $\beta (1-q) \left(\mathbf{X}^{(k-2,k-2)} - \mathbf{m}^{(k-2)} \right)$ which is ≈ 0, according to (5.1). Therefore, if we define

$$\overline{\mathbf{m}}^{(k)} \stackrel{\text{def}}{=} \mathrm{Th}\left(\mathbf{g}^{(k-1)} \mathbf{M}^{(k-1,k-2)} + \sum_{t=1}^{k-2} \gamma_t \xi^{(t)} \right),$$

then we have

$$\hat{\mathbf{m}}^{(k)} \approx \overline{\mathbf{m}}^{(k)}, \ k \leq J+1 \tag{5.10}$$

under COND (J), but also

$$\left\langle \xi^{(m)}, \hat{\mathbf{m}}^{(k)} \right\rangle \simeq \left\langle \xi^{(m)}, \overline{\mathbf{m}}^{(k)} \right\rangle \tag{5.11}$$

for $m < k \leq J+1$. This last relation follows in the same way as (5.3) in the proof of Lemma 5.1.

Lemma 5.3. *We assume* COND (J).

(a)

$$\left\langle \xi^{(s)}, \overline{\mathbf{m}}^{(J)} \right\rangle \simeq \begin{cases} \beta (1-q) \gamma_s & \text{for } s < J-1 \\ \beta (1-q) \sqrt{q - \Gamma_{J-2}^2} & \text{for } s = J-1 \end{cases}.$$

(b)

$$\left\langle \overline{\mathbf{m}}^{(J+1)}, \overline{\mathbf{m}}^{(j)} \right\rangle \simeq \rho_j,$$

for $j \leq J$, *and*

$$\left\langle \overline{\mathbf{m}}^{(J+1)}, \overline{\mathbf{m}}^{(J+1)} \right\rangle \simeq q.$$

Proof. (a) Consider first the case $s = J - 1$,

$$\overline{\mathbf{m}}^{(J)} = \mathrm{Th}\left(\left\|\mathbf{M}^{(J-1)}\right\| \xi^{(J-1)} + \sum_{t=1}^{J-2} \gamma_t \xi^{(t)}\right),$$

$$\frac{1}{N}\sum_{i=1}^{N} \xi_i^{(J-1)} \overline{m}_i^{(J)} = \frac{1}{N}\sum_{i=1}^{N} \xi_i^{(J-1)}\, \mathrm{Th}\left(\left\|\mathbf{M}^{(J-1)}\right\| \xi_i^{(J-1)} + \sum_{t=1}^{J-2} \gamma_t \xi_i^{(t)}\right).$$

We condition on $\mathcal{F}_{J-2}$. Then $\xi^{(J-1)}$ is conditionally Gaussian with covariances given in Lemma 4.2 (a), (b). We can therefore apply Lemma A.3 which gives, conditionally on $\mathcal{F}_{J-2}$, on an event $B_{J-2} \in \mathcal{F}_{J-2}$ which has probability $\geq 1 - C\exp[-N/C]$,

$$\begin{aligned}
\frac{1}{N}\sum_{i=1}^{N} \xi_i^{(J-1)} \overline{m}_i^{(J)} &\simeq \frac{1}{N}\sum_{i=1}^{N} E Z_{J-1}\, \mathrm{Th}\left(\left\|\mathbf{M}^{(J-1)}\right\| Z_{J-1} + \sum_{t=1}^{J-2} \gamma_t \xi_i^{(t)}\right) \\
&= \frac{1}{N}\sum_{i=1}^{N} \beta \left\|\mathbf{M}^{(J-1)}\right\| \left[1 - E\,\mathrm{Th}^2\left(\left\|\mathbf{M}^{(J-1)}\right\| Z_{J-1} + \sum_{t=1}^{J-2} \gamma_t \xi_i^{(t)}\right)\right] \\
&\simeq \beta\sqrt{q - \Gamma_{J-1}^2}\,\frac{1}{N}\sum_{i=1}^{N}\left[1 - E\,\mathrm{Th}^2\left(\sqrt{q - \Gamma_{J-1}^2}\, Z_{J-1} + \sum_{t=1}^{J-2} \gamma_t \xi_i^{(t)}\right)\right].
\end{aligned}$$

Applying now Lemma A.3 successively to $\xi^{(J-2)}, \xi^{(J-2)}, \ldots$, we get

$$\begin{aligned}
\frac{1}{N}\sum_{i=1}^{N} \xi_i^{(J-1)} \overline{m}_i^{(J)} &\simeq \beta\sqrt{q - \Gamma_{J-1}^2}\left[1 - E\,\mathrm{Th}^2\left(\sqrt{q - \Gamma_{J-1}^2}\, Z_{J-1} + \sum_{t=1}^{J-2} \gamma_t Z_t\right)\right] \\
&= \beta\sqrt{q - \Gamma_{J-1}^2}\,(1 - q)\,.
\end{aligned}$$

The case $s < J - 1$ uses a minor modification of the argument. One first uses Lemma A.3 successively to get

$$\frac{1}{N}\sum_{i=1}^{N} \xi_i^{(s)} \overline{m}_i^{(J)} \simeq \frac{1}{N}\sum_{i=1}^{N} \xi_i^{(s)} E\,\mathrm{Th}\left(\left\|\mathbf{M}^{(J-1)}\right\| Z_{J-1} + \sum_{t=s+1}^{J-2} \gamma_t Z_t + \gamma_s \xi_i^{(s)} + \sum_{t=1}^{s-1} \gamma_t \xi_i^{(t)}\right),$$

and then one argues as above to obtain

$$\begin{aligned}
\frac{1}{N}\sum_{i=1}^{N} \xi_i^{(s)} \overline{m}_i^{(J)} &\simeq E Z_s\, \mathrm{Th}\left(\left\|\mathbf{M}^{(J-1)}\right\| Z_{J-1} + \sum_{t=1}^{J-2} \gamma_t Z_t\right) \\
&= \beta\gamma_s\left[1 - E\,\mathrm{Th}^2\left(\left\|\mathbf{M}^{(J-1)}\right\| Z_{J-1} + \sum\nolimits_{t=1}^{J-2} \gamma_t Z_t\right)\right] = \beta\gamma_s\,(1 - q)\,.
\end{aligned}$$

(b) This is proved with a modification of the reasoning in (a).

Assume first $j \leq J$,

$$\frac{1}{N}\sum_{i=1}^{N}\overline{m}_i^{(J+1)}\overline{m}_i^{(j)} \simeq \frac{1}{N}\sum_{i=1}^{N}\left[E\,\mathrm{Th}\left(\left\|\mathbf{M}^{(J)}\right\| Z_J + \sum_{t=1}^{J-1}\gamma_t\xi_i^{(t)}\right)\right.$$
$$\left.\times\,\mathrm{Th}\left(\left\|\mathbf{M}^{(j-1)}\right\|\xi_i^{(j-1)} + \sum_{t=1}^{j-2}\gamma_t\xi_i^{(t)}\right)\right].$$

In the case $j = J+1$, the outcome is similar, one only has to replace the second factor by $\mathrm{Th}\left(\left\|\mathbf{M}^{(J)}\right\| Z_J + \sum_{t=1}^{J-1}\gamma_t\xi_i^{(t)}\right)$.

The next observation is that by the induction hypothesis, one can replace $\left\|\mathbf{M}^{(J)}\right\|$ by $\sqrt{q-\Gamma_{J-1}^2}$ and we get

$$\frac{1}{N}\sum_{i=1}^{N}\overline{m}_i^{(J+1)}\overline{m}_i^{(j)} \simeq \frac{1}{N}\sum_{i=1}^{N}\Big[E\,\mathrm{Th}\left(\sqrt{q-\Gamma_{J-1}^2}Z_J + \sum_{t=1}^{J-1}\gamma_t\xi_i^{(t)}\right)$$
$$\times\,\mathrm{Th}\left(\left\|\mathbf{M}^{(j-1)}\right\|\xi_i^{(j-1)} + \sum_{t=1}^{j-2}\gamma_t\xi_i^{(t)}\right)\Big]$$

in the $j \leq J$ case, and

$$\frac{1}{N}\sum_{i=1}^{N}\overline{m}_i^{(J+1)2} \simeq \frac{1}{N}\sum_{i=1}^{N}E\,\mathrm{Th}^2\left(\sqrt{q-\Gamma_{J-1}^2}Z_J + \sum_{t=1}^{J-1}\gamma_t\xi_i^{(t)}\right).$$

The important point is that the factor before Z_J is replaced by a constant, which is due to the induction hypothesis. We can now proceed in the same way with $\xi^{(J-1)}$, applying again Lemma A.3, conditioned on $\mathcal{F}_{J-2}$, and the induction hypothesis. The final outcome is

$$\frac{1}{N}\sum_{i=1}^{N}\overline{m}_i^{(J+1)}\overline{m}_i^{(j)} \simeq E\left[\mathrm{Th}\left(\sqrt{q-\Gamma_{J-1}^2}Z_J + \sum_{r=j}^{J-1}\gamma_r Z_r + \sum_{r=1}^{j-1}\gamma_r Z_r\right)\right.$$
$$\left.\times\,\mathrm{Th}\left(\sqrt{q-\Gamma_{j-1}^2}Z_j + \sum_{r=1}^{j-2}\gamma_r Z_r\right)\right],$$

in the case $j \leq J$, and

$$\frac{1}{N}\sum_{i=1}^{N}\overline{m}_i^{(J+1)2} \simeq E\,\mathrm{Th}^2\left(\sqrt{q-\Gamma_{J-1}^2}Z_J + \sum_{r=j}^{J-1}\gamma_r Z_r + \sum_{r=1}^{j-1}\gamma_r Z_r\right).$$

For the latter case, the right-hand side is simply q. For the case $j \leq J$, we can rewrite the expression on the right-hand side as

$$E\,\mathrm{Th}\left(\sqrt{q-\Gamma_{j-1}^2}Z'' + \gamma_{j-1}Z' + \Gamma_{j-2}Z\right)\mathrm{Th}\left(\sqrt{q-\Gamma_{j-2}^2}Z' + \Gamma_{j-2}Z\right). \quad (5.12)$$

We represent

$$\sqrt{q-\Gamma_{j-1}^2}Z''+\gamma_{j-1}Z'=aZ_1+bZ_2,$$
$$\sqrt{q-\Gamma_{j-2}^2}Z'=aZ_1+bZ_3.$$

Solving, we get $a^2+b^2=q-\Gamma_{j-2}^2$, and

$$a^2=\gamma_{j-1}\sqrt{q-\Gamma_{j-2}^2}.$$

Using this, we get that (5.12) equals

$$E\,\mathrm{Th}\left(\Gamma_{j-2}Z+aZ_1+bZ_2\right)\mathrm{Th}\left(\Gamma_{j-2}+aZ_1+bZ_2\right)=\psi\left(\Gamma_{j-2}^2+a^2\right),$$
$$\Gamma_{j-2}^2+a^2=\Gamma_{j-2}^2+\gamma_{j-1}\sqrt{q-\Gamma_{j-2}^2}=\rho_{j-1}.$$

Therefore, for $j\leq J$, we get

$$\frac{1}{N}\sum_{i=1}^N\overline{m}_i^{(J+1)}\overline{m}_i^{(j)}\simeq\psi\left(\rho_{j-1}\right)=\rho_j.$$

□

Remark 5.4. In a way, the key result of this paper is that

$$\mathbf{m}^{(k)}\approx\hat{\mathbf{m}}^{(k)} \tag{5.13}$$

holds for all k. This is correct for all β. The key point with (2.1) is that the first summand $\left\|\mathbf{M}^{(k-1)}\right\|\xi^{(k-1)}$ disappears for $k\to\infty$ as $\left\|\mathbf{M}^{(k-1)}\right\|\simeq\sqrt{q-\Gamma_{k-2}^2}$, so that for large k, $\hat{\mathbf{m}}^{(k)}$ stabilizes to $\mathrm{Th}\left(\sum_t\gamma_t\xi^{(t)}\right)$, but above the AT-line $q-\Gamma_{k-2}^2$ does not converge to 0. Therefore, above the AT-line, in every iteration, new conditionally independent contributions appear.

Acknowledgements. I thank two anonymous referees for their careful reading and useful comments which lead to a number of improvements.

A. Appendix

Lemma A.1. *Let* $\zeta=(\zeta_i)_{i=1,\ldots,N}$ *be a sequence of Gaussian vectors with* $\sup_{N,i}\mathbb{E}\left(\zeta_i^2\right)<\infty$*, and* $\sup_{N,i\neq j}N\left|\mathbb{E}\left(\zeta_i\zeta_j\right)\right|<\infty$*. Then there exist* $K,C>0$*, depending only on these two suprema, such that*

$$\mathbb{P}\left(\frac{1}{N}\sum_{i=1}^N|\zeta_i|\geq K\right)\leq C\exp\left[-N/C\right] \tag{A.1}$$

and

$$\mathbb{P}\left(\frac{1}{N}\sum_{i=1}^N\zeta_i^2\geq K\right)\leq C\exp\left[-N/C\right]. \tag{A.2}$$

Proof. We can multiply the ζ_i by a fixed positive real number. Therefore, we may assume that $\sup_{N,i\neq j} N \left|\mathbb{E}\left(\zeta_i\zeta_j\right)\right| \leq 1/4$, $\sup_{N,i} \mathbb{E}\left(\zeta_i^2\right) \leq 1$. Put $\alpha_i \overset{\text{def}}{=} 1 - \mathbb{E}\left(\zeta_i^2\right)$, and choose independent Gaussians U_i with $\mathbb{E}U_i^2 = \alpha_i$. If we prove the statements (A.1) and (A.2) for the sequence $\{\zeta_i + U_i\}$, then it follows for the ζ_i itself, simply because (A.1) and (A.2) hold for the U_i. Therefore we may assume that $\mathbb{E}\left(\zeta_i^2\right) = 1$, and $\left|\mathbb{E}\left(\zeta_i\zeta_j\right)\right| \leq 1/4N$ for $i \neq j$. Write Σ for the covariance matrix of $\{\zeta_i\}$. $\Sigma = I + \varepsilon$, where $\left|\varepsilon_{ij}\right| \leq 1/4N$. Taking the symmetric square root

$$I + \mathbf{a} = \sqrt{I + \varepsilon},$$

then $\sup_{i,j\leq N} \left|a_{ij}\right| \leq C/N$. Therefore, we can represent the ζ_i as

$$\zeta_i = Z_i + \sum_j a_{ij} Z_j,$$

where the Z_i are i.i.d. standard Gaussians. Then

$$\mathbb{P}\left(\frac{1}{N}\sum_{i=1}^N |\zeta_i| \geq K\right) \leq \mathbb{P}\left(\frac{1}{N}\sum_{i=1}^N |Z_i| \geq K/2\right) + \mathbb{P}\left(\frac{1}{N}\sum_{i=1}^N |Z_i| \geq \sqrt{K/2C}\right).$$

By choosing K appropriate, we get the desired estimate.

To prove (A.2), we use the same representation. As

$$\frac{1}{N}\sum_{i=1}^N \zeta_i^2 \leq \frac{2}{N}\sum_{i=1}^N Z_i^2 + \frac{2}{N}\sum_{i=1}^N \left(\sum_j a_{ij} Z_j\right)^2$$
$$\leq \frac{2}{N}\sum_{i=1}^N Z_i^2 + \frac{C}{N}\left(\sum_{i=1}^N |Z_i|\right)^2$$

and

$$\mathbb{P}\left(\frac{1}{N}\sum_{i=1}^N Z_i^2 \geq K\right) \leq C\exp\left[-N/C\right]$$

for large enough K, we get the desired conclusion. □

Corollary A.2. *Assume* COND (J) *and* $k \leq J$.

(a) *For any* $m \leq k$ *there exist* $C, K > 0$ *such that*

$$\mathbb{P}\left(\frac{1}{N}\sum_i \left|\xi_i^{(m)}\right| \geq K\right) \leq C\exp\left[-N/C\right].$$

(b) *For any* m, l, *there exist* $C, K > 0$ *such that*

$$\mathbb{P}\left(\frac{1}{N}\sum_i \left|\xi_i^{(m)}\xi_i^{(l)}\right| \geq K\right) \leq C\exp\left[-N/C\right].$$

(c) *If Y_i are $\mathcal{F}_{m-1}$-measurable with*

$$\mathbb{P}\left(\sup_i |Y_i| \geq K\right) \leq C \exp[-N/C]$$

for some K, then

$$\mathbb{P}\left(\left|\left\langle \xi^{(m)}, \mathbf{Y}\right\rangle\right| \geq t\right) \leq C \exp\left[-t^2 N/C\right], \quad t \leq 1.$$

Proof. Conditioned on $\mathcal{F}_{m-1}$, $\xi^{(m)}$ is Gaussian with covariances given by Lemma 4.2. On $\mathcal{F}_{m-1}$-measurable events B_N with $\mathbb{P}(B_N) \geq 1 - C\exp[-N/C]$, the variables appearing in this lemma on the right hand sides are appropriately bounded. So, on B_N, the $\xi_i^{(m)}$ are Gaussians which satisfy the conditions of the previous lemma. So (a) follows from that lemma. For (b), we estimate

$$\frac{1}{N}\sum_i \left|\xi_i^{(m)}\xi_i^{(l)}\right| \leq \sqrt{\frac{1}{N}\sum_i \xi_i^{(m)2}}\sqrt{\frac{1}{N}\sum_i \xi_i^{(l)2}},$$

so that we see that it suffices to consider $l = m$. Then we apply the lemma, part (b).

As for (c), we have that the conditional distribution of $\sqrt{N}\left\langle \xi^{(m)}, \mathbf{Y}\right\rangle$, given $\mathcal{F}_{m-1}$, is Gaussian, with bounded variance. So the statement follows. □

Lemma A.3. *Let $\left\{\eta_i^{(N)}\right\}_{i\leq N}$, be Gaussian vectors with $\sigma_{ij}^{(N)} = \mathbb{E}\eta_i^{(N)}\eta_j^{(N)}$. We assume that for some sequence $\mu_N > 0$ with $\log \mu_N$ being bounded, one has*

$$\left|\sigma_{ii}^{(N)} - \mu_N\right| \leq C/N,$$

and there are vectors $\left\{x_i^{(N)}\right\}_{i\leq N}$, $\left\{y_i^{(r,N)}\right\}_{i\leq N,\ r\leq m}$, m fixed, which are bounded in all indices, such that

$$\sup_{i\neq j,N} N^2\left|\sigma_{ij}^{(N)} - \frac{x_i^{(N)}x_j^{(N)}}{N} + \sum_{r=1}^m \frac{y_i^{(N,r)}y_j^{(N,r)}}{N}\right| < \infty.$$

Let also $F_{N,i}$, $i \leq N$, be functions $\mathbb{R} \to \mathbb{R}$, which are bounded and Lipshitz, uniformly in N, i. Then

$$\frac{1}{N}\sum_{i=1}^N F_{N,i}\left(\eta_i^{(N)}\right) \simeq \frac{1}{N}\sum_{i=1}^N E F_{N,i}\left(\sqrt{\mu_N} Z\right).$$

Proof. We leave out N in notations, as often as possible. Consider

$$\eta_i' \stackrel{\text{def}}{=} \eta_i + \sum_{r=1}^m \frac{y_i^{(r)}}{\sqrt{N}} Z_r + \sqrt{K}\frac{Z_i'}{\sqrt{N}}.$$

The constant $K > 0$ will be specified below. Then

$$\left|\frac{1}{N}\sum_{i=1}^N F_{N,i}(\eta_i) - \frac{1}{N}\sum_{i=1}^N F_{N,i}\left(\eta_i'\right)\right|$$
$$\leq Lc\sum_{r=1}^m \frac{1}{\sqrt{N}}|Z_r| + L\frac{\sqrt{K}}{N^{3/2}}\sum_{i=1}^N \left|Z_i'\right|,$$

where L is a bound on the Lipshitz constants for the $F_{N,i}$, and c is a bound of the $\left|y_i^{(r)}\right|$. As

$$P\left(|Z_r| \geq t\sqrt{N}\right) \leq C\exp\left[-t^2 N/C\right], \tag{A.3}$$

$$P\left(\frac{1}{N}\sum_{i=1}^{N}\left|Z_i'\right| \geq t\sqrt{N}\right) \leq C\exp\left[-t^2 N/C\right],$$

we get

$$\frac{1}{N}\sum_{i=1}^{N} F_{N,i}\left(\eta_i\right) \simeq \frac{1}{N}\sum_{i=1}^{N} F_{N,i}\left(\eta_i'\right).$$

$$E\left(\eta_i'^2\right) = \mu_N + \delta_i + \sum_{r=1}^{m}\frac{y_i^{(r)2}}{N} + \frac{K}{N},$$

$$E\left(\eta_i'\eta_j'\right) = \frac{x_i x_j}{N} + r_{ij}, \quad i \neq j,$$

where

$$\delta_i \stackrel{\text{def}}{=} \sigma_{ii} - \mu_N,$$

$$r_{ij} \stackrel{\text{def}}{=} \sigma_{ij} - \frac{x_i x_j}{N} + \sum_{r=1}^{m}\frac{y_i^{(r)}y_j^{(r)}}{N}.$$

We choose K large enough such that the $N \times N$-matrix Γ which is $\left(r_{ij}\right)$ off diagonal, and

$$\sum_{r=1}^{m}\frac{y_i^{(r)2}}{N} + \frac{K}{N} - \frac{x_i^2}{N} + \delta_i$$

on the diagonal is positive definite. This is possible as $\left|r_{ij}\right| \leq CN^{-2}$.

Let $\{U_i\}$ be a Gaussian vector with covariance matrix Γ. Then

$$\sqrt{\mu_N}Z_i + \frac{x_i}{\sqrt{N}}Z + U_i$$

has the same distribution as $\left\{\eta_i'\right\}$. Here we assume that $\{U_i\}$ is independent of the Z's. So, we assume that the η_i' are presented in this way,

$$\left|\frac{1}{N}\sum_{i=1}^{N} F_{N,i}\left(\eta_i'\right) - \frac{1}{N}\sum_{i=1}^{N} F_{N,i}\left(\sqrt{\mu_N}Z_i\right)\right| \leq CL\frac{|Z|}{\sqrt{N}} + L\frac{1}{N}\sum_{i=1}^{N}|U_i|.$$

We apply Lemma A.1 to the vector $\left(\sqrt{N}U_i\right)_{1\leq i\leq N}$, and (A.3) to the first summand on the right-hand side, obtaining

$$\frac{1}{N}\sum_{i=1}^{N} F_{N,i}\left(\eta_i'\right) \simeq \frac{1}{N}\sum_{i=1}^{N} F_{N,i}\left(\sqrt{\mu_N}Z_i\right),$$

so we finally have

$$\frac{1}{N}\sum_{i=1}^{N} F_{N,i}\left(\sqrt{\mu_N} Z_i\right) \simeq \frac{1}{N}\sum_{i=1}^{N} E F_{N,i}\left(\sqrt{\mu_N} Z\right),$$

by standard Gaussian isoperimetry (see e.g. [4]). □

References

1. Bayati, M., Montanari, A.: The dynamics of message passing on dense graphs, with applications to compressed sensing. IEEE Trans. Inf. Th. **57**, 771–808 (2011)
2. Bayati, M., Lelargey, M., Montanari, A.: *Universality in polytope phase transitions and message passing algorithms.* Preprint, available at http://arxiv.org/abs/1207.7321v1 [math.PR], 2012
3. de Almeida, J.R.L., Thouless, D.J.: Stability of the Sherrington–Kirkpatrick model of spin glasses. J. Phys. A. Math. Gen. **II**, 983–990 (1978)
4. Ledoux, M., Talagrand, M.: *Probability in Banach Spaces.* Berlin: Springer, 1991
5. Mézard, M., Parisi, G., Virasoro, M.: *Spin glass theory and beyond.* Singapore: World Scientific, 1987
6. Panchenko, D.: *The Sherrington–Kirkpatrick Model.* Springer Monographs in Mathematics, New York: Springer, 2013
7. Talagrand, M.: *Mean Field Models in Spin Glasses, Vol I.* Berlin: Springer, 2010
8. Talagrand, M.: The Parisi formula. Ann. Math. **163**, 221–263 (2006)
9. Thouless, D.J., Anderson, P.W., Palmer, R.G.: Solution of "solvable model in spin glasses". Phil. Mag. **35**, 593–601 (1977)

Communicated by F. Toninelli