Technical appendix to "Adaptive estimation of stationary Gaussian fields"

Nicolas Verzelen

Université Paris Sud, Laboratoire de Mathématiques, UMR 8628
Orsay Cedex F-91405
e-mail: nicolas.verzelen@math.u-psud.fr

Abstract

This is a technical appendix to "Adaptive estimation of stationary Gaussian fields" [6]. We present several proofs that have been skipped in the main paper. These proofs are organised as in Section 8 of [6].

AMS 2000 subject classifications: Primary 62H11; secondary 62M40.
Keywords and phrases: Gaussian field, Gaussian Markov random field, model selection, pseudolikelihood, oracle inequalities, Minimax rate of estimation.

1. Proof of Proposition 8.1

Proof of Proposition 8.1. First, we recall the notations introduced in [3]. Let N be a positive integer. Then, $\mathcal{I}_N$ stands for the family of subsets of $\{1,\ldots,N\}$ of size less than 2. Let $\mathcal{T}$ be a set of vectors indexed by $\mathcal{I}_N$. In the sequel, $\mathcal{T}$ is assumed to be a compact subset of $\mathbb{R}^{(N(N+1)/2)+1}$. The following lemma states a slightly modified version of the upper bound in remark 7 in [3].

Lemma 1.1. *Let T be a supremum of Rademacher chaos indexed by $\mathcal{I}_N$ of the form*

$$T := \sup_{t\in\mathcal{T}}\left|\sum_{\{i,j\}} U_iU_jt_{\{i,j\}} + \sum_{i=1}^N t_{\{i\}} + t_{\varnothing}\right| ,$$

where $U_1,\ldots,U_N$ are independent Rademacher random variables. Then for any $x>0$,

$$\mathbb{P}\left\{T\geq \mathbb{E}[T]+x\right\}\leq 4\exp\left(-\frac{x^2}{L_1\mathbb{E}[D]^2}\wedge\frac{x}{L_2E}\right) , \tag{1}$$

where D and E are defined by:

$$\begin{aligned} D &:= \sup_{t\in\mathcal{T}}\sup_{\alpha:\|\alpha\|_2\leq 1}\left|\sum_{i=1}^N U_i\sum_{j\neq i}\alpha_jt_{\{i,j\}}\right| ,\\ E &:= \sup_{t\in\mathcal{T}}\sup_{\alpha^{(1)},\alpha^{(2)},\|\alpha^{(1)}\|_2\leq 1\ \|\alpha^{(2)}\|\leq 1}\left|\sum_{i=1}^N\sum_{j\neq i}t_{\{i,j\}}\alpha_i^{(1)}\alpha_j^{(2)}\right| .\end{aligned}$$

Contrary to the original result of [3], the chaos are not assumed to be homogeneous. Besides, the $t_{\{i\}}$ are redundant with $t_{\varnothing}$. In fact, we introduced this family in order to emphasize the connection with Gaussian chaos in the next result.

A suitable application of the central limit theorem enables to obtain a corresponding bound for Gaussian chaos of order 2.

Lemma 1.2. *Let T be a supremum of Gaussian chaos of order 2.*

$$T := \sup_{t \in \mathcal{T}} \left| \sum_{\{i,j\}} t_{\{i,j\}} Y_i Y_j + \sum_i t_i Y_i^2 + t_{\varnothing} \right| , \tag{2}$$

where $Y_1, \ldots, Y_N$ are independent standard Gaussian random variable. Then, for any $x > 0$,

$$\mathbb{P}\left\{T \geq \mathbb{E}[T] + x\right\} \leq \exp\left(-\frac{x^2}{\mathbb{E}[D]^2 L_1} \wedge \frac{x}{E L_2}\right) , \tag{3}$$

where

$$\begin{aligned} D &:= \sup_{t \in \mathcal{T}} \sup_{\alpha \in \mathbb{R}^N \|\alpha\|_2 \leq 1} \sum_{i,j} Y_i (1 + \delta_{i,j}) \alpha_j t_{\{i,j\}} , \\ E &:= \sup_{t \in \mathcal{T}} \sup_{\alpha_1,\ \|\alpha_1\|_2 \leq 1} \sup_{\alpha_2,\ \|\alpha_2\|_2 \leq 1} \sum_{i,j} \alpha_{1,i} \alpha_{2,j} t_{\{i,j\}} (1 + \delta_{i,j}) . \end{aligned}$$

The proof of this Lemma is postponed to the end of this section. To conclude, we derive the result of Proposition 8.1 from this last lemma. For any matrix $R \in F$, we define the vector $t^R \in \mathbb{R}^{nr(nr+1)/2+1}$ indexed by $\mathcal{I}_{nr}$ as follows

$$t^R_{\{(i,k),(j,l)\}} := \delta_{k,l}(2 - \delta_{i,j}) \frac{R[i,j]}{n}, \quad t^R_{\{(i,k)\}} := \frac{R[i,i]}{n}, \quad \text{and } t^R_{\varnothing} := -tr(R) ,$$

where $\delta_{i,j}$ is the indicator function of $i = j$. In order to apply Lemma 1.2 with $N = nr$ and $\mathcal{T} = \left\{t^R | R \in F\right\}$, we have to work out the quantities D and E.

$$\begin{aligned} D &= \sup_{t^R \in \mathcal{T}} \sup_{\alpha \in \mathbb{R}^{nr},\ \|\alpha\|_2 \leq 1} \left\{ \sum_{i=1}^r \sum_{k=1}^n Y_{[i,k]} \sum_{j=1}^r \sum_{l=1}^n t_{ij}^{R,k,l} (1 + \delta_{i,j}\delta_{k,l}) \alpha_j^l \right\} \\ &= \sup_{R \in F} \sup_{\alpha \in \mathbb{R}^{nr},\ \|\alpha\|_2 \leq 1} 2 \left\{ \sum_{i=1}^r \sum_{k=1}^n Y_{[i,k]} \sum_{j=1}^r \frac{R[i,j] \alpha_j^k}{n} \right\} \\ &= \sup_{R \in F} \sup_{\alpha \in \mathbb{R}^{nr},\ \|\alpha\|_2 \leq 1} \frac{2}{n} \left\{ \sum_{k=1}^n \sum_{j=1}^r \alpha_j^k \left(\sum_{i=1}^r Y_{[i,k]} R[i,j] \right) \right\} . \end{aligned}$$

Applying Cauchy-Schwarz identity yields

$$D^2 = \frac{4}{n^2}\sup_{R\in F}\left\{\sum_{k=1}^{n}\sum_{j=1}^{r}\left(\sum_{i=1}^{r}Y_{[i,k]}R_{[i,j]}\right)^2\right\}$$
$$= \frac{4}{n}\sup_{R\in F} tr(R\overline{YY^*}R^*) \ . \quad (4)$$

Let us now turn the constant E

$$E = \sup_{t^R\in\mathcal{T}} \sup_{\substack{\alpha_1,\alpha_2\in\mathbb{R}^{nr}\\ \|\alpha_1\|_2\leq 1, \|\alpha_2\|_2\leq 1}} \sum_{1\leq i,j\leq r}\sum_{1\leq k,l\leq n}(1+\delta_{ij}\delta_{k,l})t^{R,kl}_{i,j}\alpha^k_{1,i}\alpha^l_{2,j}$$
$$= \sup_{R\in F} \sup_{\substack{\alpha_1,\alpha_2\in\mathbb{R}^{nr}\\ \|\alpha_1\|_2\leq 1, \|\alpha_2\|_2\leq 1}} \frac{2}{n}\sum_{1\leq i,j\leq r}\sum_{1\leq k\leq n}R_{[i,j]}\alpha^k_{1,i}\alpha^k_{2,j} \ .$$

From this last expression, it follows that E is a supremum of L_2 operator norms

$$E = \frac{2}{n}\sup_{R\in F}\varphi_{\max}\left(Diag^{(n)}(R)\right) \ ,$$

where $Diag^{(n)}(R)$ is the $(nr\times nr)$ block diagonal matrix such that each diagonal block is made of the matrix R. Since the largest eigenvalue of $Diag^{(n)}(R)$ is exactly the largest eigenvalue of R, we get

$$E = \frac{2}{n}\sup_{R\in F}\varphi_{\max}(R) \ . \quad (5)$$

Applying Proposition 1.2 and gathering identities (4) and (5) yields

$$\mathbb{P}(Z\geq \mathbb{E}(Z)+t)\leq \exp\left[-\left(\frac{t^2}{L_1\mathbb{E}(V)}\bigwedge\frac{t}{L_2B}\right)\right] \ ,$$

where $B = E$ and $V = D^2$. □

Proof of Lemma 1.1. This result is an extension of Corollary 4 in [3]. We shall closely follow the sketch of their proof adapting a few arguments. First, we upper bound the moments of $(T-\mathbb{E}(T))_+$. Then, we derive the deviation inequality from it. Here, $x_+ = \max(x, 0)$.

Lemma 1.3. *For all real numbers $q\geq 2$,*

$$\|(T-\mathbb{E}(T))_+\|_q \leq \sqrt{Lq}\mathbb{E}(D) + LqE \ , \quad (6)$$

where $\|T\|_q^q$ stands for the q-th moment of the random variable T. The quantities D and E are defined in Lemma 1.1.

By Lemma 1.3, for any $t \geq 0$ and any $q \geq 2$,

$$\begin{aligned} \mathbb{P}\left(T \geq \mathbb{E}(T) + t\right) &\leq \frac{\mathbb{E}\left[(T - \mathbb{E}(T))_+^q\right]}{t^q} \\ &\leq \left(\frac{\sqrt{Lq}\mathbb{E}(D) + LqE}{t}\right)^q . \end{aligned}$$

The right-hand side is at most 2^{-q} if $\sqrt{Lq}\mathbb{E}(D) \leq t/4$ and $LqE \leq t/4$. Let us set

$$q_0 := \frac{t^2}{16L\mathbb{E}(D)^2} \wedge \frac{t}{4LE} .$$

If $q_0 \geq 2$, then $\mathbb{P}\left(T \geq \mathbb{E}(T) + t\right) \leq 2^{-q_0}$. On the other hand if $q_0 < 2$, then $4 \times 2^{-q_0} \geq 1$. It follows that

$$\mathbb{P}\left(T \geq \mathbb{E}(T) + t\right) \leq 4 \exp\left(-\frac{\log(2)}{4L}\left[\frac{t^2}{4\mathbb{E}(D)^2} \wedge \frac{t}{E}\right]\right) .$$

□

Proof of Lemma 1.3. This result is based on the entropy method developed in [3]. Let $f : \mathbb{R}^N \to \mathbb{R}$ be a measurable function such that $T = f(U_1, \ldots, U_N)$. In the sequel, $U'_1, \ldots, U'_N$ denote independent copies of $U_1, \ldots, U_N$. The random variable T'_i and V^+ are defined by

$$\begin{aligned} T'_i &:= f(U_1, \ldots, U_{i-1}, U'_i, U_{i+1}, \ldots, U_N) , \\ V^+ &:= \mathbb{E}\left[\sum_{i=1}^N (T - T'_i)_+^2 | U_1^N\right] , \end{aligned}$$

where U_1^N refers to the set $\{U_1, \ldots, U_N\}$. Theorem 2 in [3] states that for any real $q \geq 2$,

$$\|(T - \mathbb{E}(T))_+\|_q \leq \sqrt{Lq}\|\sqrt{V^+}\|_q . \tag{7}$$

To conclude, we only have bound the moments of $\sqrt{V^+}$. By definition,

$$T = \sup_{t \in \mathcal{T}} \left| \sum_{\{i,j\}} U_i U_j t_{\{i,j\}} + \sum_{i=1}^N t_{\{i\}} + t_{\varnothing} \right| .$$

Since the set $\mathcal{T}$ is compact, this supremum is achieved almost surely at an element t^0 of $\mathcal{T}$. For any $1 \leq i \leq N$,

$$(T - T'_i)_+^2 \leq \left((U_i - U'_i)\left|\sum_{j \neq i} U_j t^0\{i,j\}\right|\right)^2 .$$

Gathering this bound for any i between 1 and N, we get

$$\begin{aligned} V^{+} &\leq \sum_{i=1}^{N}\mathbb{E}\left[\left((U_i-U_i')\left|\sum_{j\neq i}U_jt^0\{i,j\}\right|\right)^2\middle|U_1^N\right] \\ &\leq 2\sum_{i=1}^{N}\left[\sum_{j\neq i}U_jt^0\{i,j\}\right]^2 \\ &\leq 2\sup_{\alpha\in\mathbb{R}^N,\ \|\alpha\|_2\leq 1}\left[\sum_{i=1}^{N}\alpha_i\left(\sum_{j\neq i}t^0_{\{i,j\}}U_j\right)\right]^2 \\ &\leq 2\sup_{t\in\mathcal{T}}\sup_{\alpha\in\mathbb{R}^N,\ \|\alpha\|_2\leq 1}\sum_{i=1}^{N}\left[U_i\sum_{j\neq i}\alpha_jt_{\{i,j\}}\right]^2 = 2D^2\ . \end{aligned}$$

Combining this last bound with (7) yields

$$\begin{aligned} \|(T-\mathbb{E}(T))_+\|_q &\leq \sqrt{Lq}\sqrt{2}\|D\|_q \\ &\leq \sqrt{Lq}\left[\mathbb{E}(D)+|(D-\mathbb{E}(D))_+\|_q\right]\ . \end{aligned} \quad (8)$$

Since the random variable D defined in Lemma 1.1 is a measurable function f_2 of the variables $U_1,\ldots,U_N$, we apply again Theorem 2 in [3].

$$\|(D-\mathbb{E}(D))_+\|_q \leq \sqrt{Lq}\left\|\sqrt{V_2^+}\right\|_q ,$$

where V_2^+ is defined by

$$V_2^+ := \mathbb{E}\left[\sum_{i=1}^{N}(D-D_i')^2_+\middle|U_i^N\right],$$

and $D_i' := f_2(U_1,\ldots,U_{i-1},U_i',U_{i+1},\ldots,U_N)$. As previously, the supremum in D is achieved at some random parameter (t^0,α^0). We therefore upper bound V_2^+ as previously.

$$\begin{aligned} V_2^{+} &\leq \sum_{i=1}^{N}\mathbb{E}\left[\left((U_i-U_i')\left(\sum_{j\neq i}\alpha_j^0t^0_{\{i,j\}}\right)\right)^2\middle|U_1^N\right] \\ &\leq 2\sum_{i=1}^{N}\left(\sum_{j\neq i}\alpha_j^0t^0_{\{i,j\}}\right)^2 \\ &\leq 2\sup_{\alpha^{(2)}\in\mathbb{R}^N,\|\alpha\|_2\leq 1}\left(\sum_{i=1}^{N}\alpha_j^{(2)}\sum_{j\neq i}\alpha_i^0t_{\{i,j\}}\right)^2 = 2E^2\ . \end{aligned}$$

Gathering this upper bound with (8) yields

$$\|(T-\mathbb{E}(T))_+\|_q \leq \sqrt{Lq}\mathbb{E}(D)+LqE\ .$$

□

Proof of Lemma 1.2. We shall apply the central limit theorem in order to transfer results for Rademacher chaos to Gaussian chaos. Let f be the unique function satisfying $T = f(y_1, \ldots, y_N)$ for any $(y_1, \ldots, y_N) \in \mathbb{R}^N$. As the set $\mathcal{T}$ is compact, the function f is known to be continuous. Let $(U_i^{(j)})_{1\leq i\leq N, j\geq 0}$ an i.i.d. family of Rademacher variables. For any integer $n > 0$, the random variables $Y^{(n)}$ and $T^{(n)}$ are defined by

$$\begin{aligned} Y^{(n)} &:= \left(\sum_{j=1}^{n} \frac{U_1^{(j)}}{\sqrt{n}}, \ldots, \sum_{j=1}^{n} \frac{U_N^{(j)}}{\sqrt{n}}\right) , \\ T^{(n)} &:= f\left(Y^{(n)}\right) . \end{aligned}$$

Clearly, $T^{(n)}$ is a supremum of Rademacher chaos of order 2 with nN variables and a constant term. By the central limit theorem, $T^{(n)}$ converges in distribution towards T as n tends to infinity. Consequently, deviation inequalities for the variables $T^{(n)}$ transfer to T as long as the quantities $\mathbb{E}\left[D^{(n)}\right]$, $E^{(n)}$, and $\mathbb{E}[T^{(n)}]$ converge.

We first prove that the sequence $T^{(n)}$ converges in expectation towards T. As $T^{(n)}$ converges in distribution, it is sufficient to show that the sequence $T^{(n)}$ is asymptotically uniformly integrable. The set $\mathcal{T}$ is compact, thus there exists a positive number t_∞ such that

$$\begin{aligned} T^{(n)} &\leq t_\infty\left[\sum_{i,j} |Y_i^{(n)} Y_j^{(n)}| + 1\right] \\ &\leq t_\infty\left[1 + (N+1)/2 \sum_{i=1}^{N} \left(Y_i^{(n)}\right)^2\right] . \end{aligned}$$

It follows that

$$\left(T^{(n)}\right)^2 \leq t_\infty^2 \left(\frac{N+1}{2}\right)^2 \frac{N+2}{2}\left[1 + \sum_{i=1}^{N} \left(Y_i^{(n)}\right)^4\right] . \tag{9}$$

The sequence $Y_i^{(n)}$ does not only converge in distribution to a standard normal distribution but also in moments (see for instance [1] p.391). It follows that $\overline{\lim}\mathbb{E}\left[\left(T^{(n)}\right)^2\right] \leq \infty$ and the sequence $f\left(Y^{(n)}\right)$ is asymptotically uniformly integrable. As a consequence,

$$\lim_{n\to\infty} \mathbb{E}\left[T^{(n)}\right] = \mathbb{E}[T] .$$

Let us turn to the limit of $\mathbb{E}\left[D^{(n)}\right]$. As the variable $T^{(n)}$ equals

$$T^{(n)} = \sup_{t\in\mathcal{T}} \left| \sum_{\{i,j\}} t_{\{i,j\}} \sum_{1\leq k,l\leq n} \frac{U_i^{(k)} U_j^{(l)}}{n} + \sum_i t_i \sum_{1\leq k\leq n} \frac{U_i^{(k)}}{\sqrt{n}} \sum_{l\neq k} \frac{U_i^{(l)}}{\sqrt{n}} + t_\varnothing + \sum_i t_i \right| ,$$

it follows that

$$\begin{aligned} D^{(n)} &= \sup_{t\in\mathcal{T}} \sup_{\alpha\in\mathbb{R}^{nN},\ \|\alpha\|_2\leq 1} \Big| \sum_{1\leq i\leq N}\sum_{1\leq k\leq n} U_i^{(k)} \Big\{ \sum_{j\neq i} \frac{t_{\{i,j\}}}{n} \sum_{1\leq l\leq n} \alpha_j^{(l)} + 2\sum_{l\neq k} 2\frac{t_{\{i\}}}{n}\alpha_i^{(l)} \Big\} \Big| \\ &\leq \sup_{t\in\mathcal{T}} \sup_{\alpha\in\mathbb{R}^{nN},\ \|\alpha\|_2\leq 1} \Big\{ \sum_i \frac{U_i^{(k)}}{\sqrt{n}} \sum_j (1+\delta_{i,j}) t_{\{i,j\}} \frac{\sum_{1\leq l\leq n}\alpha_j^{(l)}}{\sqrt{n}} \Big\} + A^{(n)} , \end{aligned} \quad (10)$$

where the random variable $A^{(n)}$ is defined by

$$A^{(n)} := \sup_{t\in\mathcal{T}} \sup_{\alpha\in\mathbb{R}^{nN},\ \|\alpha\|_2\leq 1} \sum_{i=1}^N \sum_{j=1}^n t_{\{i\}} \frac{U_i^{(j)}}{n} \alpha_i^j .$$

Straightforwardly, one upper bounds $A^{(n)}$ by $t_\infty/n\sqrt{\sum_{i=1}^N\sum_{j=1}^n \left(U_i^{(j)}\right)^2}$ and its expectation satisfies

$$\mathbb{E}\left(\left|A^{(n)}\right|\right) \leq t_\infty\sqrt{\frac{N}{n}} ,$$

which goes to 0 when n goes to infinity. Thus, we only have to upper bound the expectation of the first term in (10). Clearly, the supremum is achieved only when for all $1\leq j\leq N$, the sequence $(\alpha_j^{(l)})_{1\leq l\leq n}$ is constant. In such a case, the sequence $(\alpha_j^{(1)})_{1\leq j\leq N}$ satisfies $\|\alpha^{(1)}\|_2 \leq 1/\sqrt{n}$. it follows that

$$\mathbb{E}\left[D^{(n)}\right] = \mathbb{E}\Big\{ \sup_{t\in\mathcal{T}} \sup_{\alpha\in\mathbb{R}^N \|\alpha\|_2\leq 1} \mathbb{E}\Big[\sum_i Y_i^{(n)} \sum_j (1+\delta_{i,j})\alpha_j \Big] \Big\} + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right) .$$

Let g be the function defined by

$$g\left(y_1,\ldots,y_N\right) = \sup_{t\in\mathcal{T}} \sup_{\alpha\in\mathbb{R}^N \|\alpha\|_2\leq 1} \Big[\sum_i y_i \sum_j (1+\delta_{i,j})\alpha_j \Big] ,$$

for any $(y_1,\ldots,y_N)\in\mathbb{R}^N$. The function $g(.)$ is measurable and continuous as the supremum is taken over a compact set. As a consequence, $g(Y^{(n)})$ converges in distribution towards $g(Y)$. As previously, the sequence is asymptotically uniformly integrable since its moment of order 2 is uniformly upper bounded. It follows that $\lim \mathbb{E}\left[D^{(n)}\right] = \mathbb{E}\left[D\right]$.

Third, we compute the limit of $E^{(n)}$. By definition,

$$\begin{aligned} E^{(n)} &= \sup_{t\in\mathcal{T}} \sup_{\alpha_1,\alpha_2\in\mathbb{R}^{nN},\ \|\alpha_1\|_2\leq 1, \|\alpha_2\|_2\leq 1} \sum_{i=1}^N\sum_{k=1}^n \alpha_{1,i}^k \Big[\sum_{j\neq i}\sum_{l=1}^n \alpha_{2,j}^{(l)} \frac{t_{\{i,j\}}}{n} + 2\sum_{l\neq k} \alpha_{2,i}^{(l)} \frac{t_{\{i\}}}{n} \Big] \\ &= \sup_{t\in\mathcal{T}} \sup_{\alpha_1,\alpha_2,\ \|\alpha_1\|_2\leq 1, \|\alpha_2\|_2\leq 1} \sum_{i=1}^N\sum_{j=1}^N (1+\delta_{i,j}) \frac{t_{\{i,j\}}}{n} \Big[\sum_{k=1}^n\sum_{l=1}^n \alpha_{1,i}^{(k)}\alpha_{2,j}^{(l)} \Big] + \mathcal{O}\left(\frac{1}{n}\right) . \end{aligned}$$

As for the computation of $D^{(n)}$, the supremum is achieved when the sequences $(\alpha^k_{1,i})_{1\leq k\leq n}$ and $(\alpha^l_{2,j})_{1\leq l\leq n}$ are constant for any $i\in\{1,\ldots,N\}$. Thus, we only have to consider the supremum over the vectors α_1 and α_2 in $\mathbb{R}^N$.

$$E^{(n)} = \sup_{t\in\mathcal{T}}\sup_{\alpha_1,\alpha_2\in\mathbb{R}^N\|\alpha_i\|_2\leq 1}\sum_{i=1}^N\sum_{j=1}^N(1+\delta_{ij})t_{i,j}\alpha_{1,i}\alpha_{2,j}+\mathcal{O}\left(\frac{1}{n}\right).$$

It follows that $E^{(n)}$ converges towards E when n tends to infinity.

The random variable $T^{(n)}-\mathbb{E}(T^{(n)})$ converges in distribution towards $T-\mathbb{E}(T)$. By Lemma 1.1 ,

$$\mathbb{P}(T-\mathbb{E}(T)\geq x)\leq \underline{\lim}\ \exp\left(-\frac{x^2}{\mathbb{E}[D^{(n)}]^2L_1}\wedge\frac{x}{E^{(n)}L_2}\right)\ ,$$

for any $x>0$. Combining this upper bound with the convergence of the sequences $D^{(n)}$ and $E^{(n)}$ allows to conclude. □

2. Proof of Theorem 3.1

Proof of Lemma 8.3. We only consider here the anisotropic case, since the isotropic case is analogous. This result is based on the deviation inequality for suprema of Gaussian chaos of order 2 stated in Proposition 8.1. For any model m' belonging to $\mathcal{M}$, we shall upper bound the quantities $\mathbb{E}(Z_{m'})$, $B_{m'}$, and $\mathbb{E}(W_{m'})$ defined in (42) in [6].

1. Let us first consider the expectation of $Z_{m'}$. Let $U'_{m,m'}$ be the new vector space defined by

$$U'_{m,m'} := U_{m,m'}\frac{\sqrt{D_\Sigma}}{p}\ ,$$

where $U_{m,m'}$ is introduced in the proof of Lemma 8.2 in [6]. This new space allows to handle the computation with the canonical inner product in the space of matrices. Let $\mathcal{B}^{(2)}_{m^2,m'^2}$ be the unit ball of $U'_{m,m'}$ with respect to the canonical inner product. If R belongs to $U_{m,m'}$, then $\|R\|_{\mathcal{H}'}=\|R\sqrt{D_\Sigma}/p\|_F$, where $\|.\|_F$ stands for the Frobenius norm.

$$\begin{aligned}Z_{m'} &= \sup_{R\in\mathcal{B}^{\mathcal{H}'}_{m^2,m'^2}}\frac{1}{p^2}tr\left[RD_\Sigma(\overline{\mathbf{YY}^*}-I_{p^2})\right]\\ &= \sup_{R\in\mathcal{B}^{(2)}_{m^2,m'^2}} tr\left[R\frac{\sqrt{D_\Sigma}}{p}\left(\overline{\mathbf{YY}^*}-I_{p^2}\right)\right] \qquad (11)\\ &= \left\|\Pi_{U'_{m,m'}}\frac{\sqrt{D_\Sigma}}{p}\left(\overline{\mathbf{YY}^*}-I_{p^2}\right)\right\|_F\ ,\end{aligned}$$

where $\Pi_{U'_{m,m'}}$ refers to the orthogonal projection with respect to the canonical inner product onto the space $U'_{m,m'}$. Let $F_1, \ldots, F_{d_{m^2,m'^2}}$ denote an orthonormal basis of $U'_{m,m'}$.

$$\begin{aligned}
\mathbb{E}(Z^2_{m'}) &= \sum_{i=1}^{d_{m^2,m'^2}} \mathbb{E}\left[tr^2\left(F_i\sqrt{\frac{D_\Sigma}{p^2}}\left(\overline{\mathbf{Y}\mathbf{Y}^*} - I_{p^2}\right)\right)\right] \\
&= \sum_{i=1}^{d_{m^2,m'^2}} \mathbb{E}\left[\sum_{j=1}^{p^2} F_{i[j,j]}\frac{\sqrt{D_{\Sigma[j,j]}}}{p}(\overline{\mathbf{Y}\mathbf{Y}^*}_{[j,j]} - 1)\right]^2 \\
&= \sum_{i=1}^{d_{m^2,m'^2}} \frac{2}{np^2} tr(F_i D_\Sigma F_i) \\
&\leq \sum_{i=1}^{d_{m^2,m'^2}} \frac{2\varphi_{\max}(D_\Sigma)}{np^2} = \frac{2d_{m^2,m'^2}\varphi_{\max}(\Sigma)}{np^2} .
\end{aligned}$$

Applying Cauchy-Schwarz inequality, it follows that

$$\mathbb{E}(Z_{m'}) \leq \sqrt{\frac{2d_{m^2,m'^2}\varphi_{\max}(\Sigma)}{np^2}} . \tag{12}$$

2. Using the identity (11), the quantity $B_{m'}$ equals

$$B_{m'} = \frac{2}{n} \sup_{R \in \mathcal{B}^{(2)}_{m^2,m'^2}} \varphi_{\max}\left(R\frac{\sqrt{D_\Sigma}}{p}\right) .$$

As the operator norm is under-multiplicative and as it dominates the Frobenius norm, we get the following bound

$$B_{m'} \leq \frac{2\sqrt{\varphi_{\max}(\Sigma)}}{np} . \tag{13}$$

3. Let us turn to bounding the quantity $\mathbb{E}(W_{m'})$. Again, by introducing the ball $\mathcal{B}^{(2)}_{m^2,m'^2}$, we get

$$\begin{aligned}
W_{m'} &= \frac{4}{n} \sup_{R \in \mathcal{B}^{\mathcal{H}'}_{m^2,m'^2}} \frac{1}{p^2} tr\left[R\overline{\mathbf{Y}\mathbf{Y}^*}D_\Sigma R\right] \\
&\leq \frac{4\varphi_{\max}(\Sigma)}{np^2} \sup_{R \in \mathcal{B}^{(2)}_{m^2,m'^2}} tr\left[R\overline{\mathbf{Y}\mathbf{Y}^*}R\right] \\
&\leq \frac{4\varphi_{\max}(\Sigma)}{np^2}\left(1 + \sup_{R \in \mathcal{B}^{(2)}_{m^2,m'^2}} tr\left[R\left(\overline{\mathbf{Y}\mathbf{Y}^*} - I_{p^2}\right)R\right]\right) .
\end{aligned}$$

Let $F_1, \ldots F_{d_{m^2,m'^2}}$ an orthonormal basis of $U'_{m,m'}$ and let λ be a vector in $\mathbb{R}^{d_{m^2,m'^2}}$. We write $\|\lambda\|_2$ for its L_2 norm.

$$\begin{aligned}
&\mathbb{E}\left(\sup_{R\in\mathcal{B}^{(2)}_{m^2,m'^2}} tr\left[R\left(\overline{\mathbf{YY}^*}-I_{p^2}\right)R\right]^2\right)\\
&\quad= \mathbb{E}\left(\sup_{\|\lambda\|_2\leq 1}\sum_{i,j=1}^{d_{m^2,m'^2}}\lambda_i\lambda_j tr\left[F_iF_j(\overline{\mathbf{YY}^*}/n-I_{p^2})\right]\right)^2\\
&\quad\leq \sum_{i,j=1}^{d_{m^2,m'^2}}\mathbb{E}\left(tr\left[F_iF_j(\mathbf{YY}^*/n-I_{p^2})\right]^2\right)\ .
\end{aligned}$$

The second inequality is a consequence of Cauchy-Schwarz inequality in $\mathbb{R}^{(d_{m^2,m'^2})^2}$ since the l_2 norm of the vector $(\lambda_i\lambda_j)_{1\leq i,j\leq d_{m^2,m'^2}}\in\mathbb{R}^{d^2_{m^2,m'^2}}$ is bounded by 1. Since the matrices F_i are diagonal, we get

$$\mathbb{E}\left(\sup_{R\in\mathcal{B}^{(2)}_{m^2,m'^2}} tr\left[R(\mathbf{YY}^*/n-I)R\right]^2\right)\leq\frac{2}{n}\sum_{i,j=1}^{d_{m^2,m'^2}}\|F_iF_j\|_2^2\ .$$

It remains to bound the norm of the products F_iF_j for any i,j between 1 and d_{m^2,m'^2}.

$$\sum_{i,j=1}^{d_{m^2,m'^2}}\|F_iF_j\|_2^2 = \sum_{i,j=1}^{d_{m^2,m'^2}}\sum_{k=1}^{p^2}F_{i[k,k]}{}^2F_{j[k,k]}{}^2=\sum_{k=1}^{p^2}\left(\sum_{i=1}^{d_{m^2,m'^2}}F_{i[k,k]}{}^2\right)^2\ .$$

For any $k\in\{1,\ldots,p^2\}$, $\sum_{i=1}^{d_{m^2,m'^2}}F_{i[k,k]}{}^2\leq 1$ since $(F_1,\ldots,F_{d_{m^2,m'^2}})$ form an orthonormal family. Hence, we get

$$\sum_{i,j=1}^{d_{m^2,m'^2}}\|F_iF_j\|_2^2 \leq \sum_{k=1}^{p^2}\sum_{i=1}^{d_{m^2,m'^2}}F_{i[k,k]}{}^2=d_{m^2,m'^2}\ .$$

All in all, we have proved that

$$\mathbb{E}(W_{m'})\leq\frac{4\varphi_{\max}(\Sigma)}{np^2}\left[1+\sqrt{\frac{2d_{m^2,m'^2}}{n}}\right]. \tag{14}$$

Gathering these three bounds and applying Proposition 8.1 allows to obtain the following deviation inequality:

$$\begin{aligned}
&\mathbb{P}\left(Z_{m'}\geq\sqrt{\tfrac{2\varphi_{\max}(\Sigma)}{n}}\left\{\sqrt{1+\alpha/2}\sqrt{d_{m^2,m'^2}}+\xi\right\}\right)\\
&\leq\exp\left\{-\left[\frac{\left[\left(\sqrt{1+\alpha/2}-1\right)\sqrt{d_{m^2,m'^2}}+\xi\right]^2}{2L_1\left(1+\sqrt{2d_{m^2,m'^2}/n}\right)}\bigwedge\frac{\sqrt{n}\left[\left(\sqrt{1+\alpha/2}-1\right)\sqrt{d_{m^2,m'^2}}+\xi\right]}{\sqrt{2}L_2}\right]\right\}\\
&\leq\exp\left\{-\left[\frac{\omega^2_{m,m'}}{2L_1\left(1+\sqrt{2d_{m^2,m'^2}/n}\right)}\bigwedge\frac{\sqrt{n}\omega_{m,m'}}{\sqrt{2}L_2}\right]-\left[\frac{\xi\omega_{m,m'}}{L_1\left[1+\sqrt{2d_{m^2,m'^2}/n}\right]}\bigwedge\frac{\sqrt{n}\xi}{\sqrt{2}L_2}\right]\right\}\ ,
\end{aligned}$$

where $\omega_{m,m'} = \left(\sqrt{1+\alpha/2}-1\right)\sqrt{d_{m^2,m'^2}}$. As n and d_{m^2,m'^2} are larger than one, there exists a universal constant L'_2 such that

$$\left[\frac{(\sqrt{1+\alpha/2}-1)^2 d_{m^2,m'^2}}{2L_1\left(1+\sqrt{2d_{m^2,m'^2}/n}\right)}\bigwedge\frac{\sqrt{n}(\sqrt{1+\alpha/2}-1)\sqrt{d_{m^2,m'^2}}}{\sqrt{2}L_2}\right] \geq 4L'_2\sqrt{d_{m^2,m'^2}}\left[\left(\sqrt{1+\alpha/2}-1\right)^2\wedge\left(\sqrt{1+\alpha/2}-1\right)\right] .$$

Since the vector space $U_{m,m'}$ contains all the matrices $D(\theta')$ with θ' belonging to m', d_{m^2,m'^2} is larger than $d_{m'}$. Besides, by concavity of the square root function, it holds that $\sqrt{1+\alpha/2}-1\geq\alpha[4\sqrt{1+\alpha/2}]^{-1}$. Setting $L'_1 := [4L_1(1+\sqrt{2})]^{-1}\wedge[\sqrt{2}L_2]^{-1}$ and arguing as previously leads to

$$\frac{\xi(\sqrt{1+\alpha/2}-1)\sqrt{d_{m^2,m'^2}}}{L_1\left(1+\sqrt{2d_{m^2,m'^2}/n}\right)}\bigwedge\frac{\sqrt{n}\xi}{\sqrt{2}L_2}\geq L'_1\xi\left[\frac{\alpha}{\sqrt{1+\alpha/2}}\wedge\sqrt{n}\right] .$$

Gathering these two inequalities allows us to conclude that

$$\mathbb{P}\left(Z_{m'}\geq\sqrt{\frac{2\varphi_{\max}(\Sigma)}{n}}\left\{\sqrt{(1+\alpha/2)\,d_{m^2,m'^2}}+\xi\right\}\right) \leq \exp\left\{-L'_2\sqrt{d_{m'}}\left(\frac{\alpha}{\sqrt{1+\alpha/2}}\wedge\frac{\alpha^2}{1+\alpha/2}\right)-L'_1\xi\left[\frac{\alpha}{\sqrt{1+\alpha/2}}\wedge\sqrt{n}\right]\right\} .$$

□

Proof of Lemma 8.4 in [6]. The approach falls in two parts. First, we relate the dimensions d_m and d_{m^2} to the number of nodes of the torus Λ that are closer than r_m or $2r_m$ to the origin $(0,0)$. We recall that the quantity r_m is introduced in Definition 2.1 of [6]. Second, we compute a nonasymptotic upper bound of the number of points in $\mathbb{Z}^2$ that lie in the disc of radius r. This second step is quite tedious and will only give the main arguments.

Let m be a model of the collection $\mathcal{M}_1$. By definition, m is the set of points lying in the disc of radius r_m centered on $(0,0)$. Hence,

$$\Theta_m = vect\left\{\Psi_{i,j},\ (i,j)\in m\right\} ,$$

where the matrices $\Psi_{i,j}$ are defined by Eq. (14) in [6]. As $\Psi_{i,j}=\Psi_{-i,-j}$, the dimension d_m of Θ_m is exactly the number of orbits of m under the action of the central symmetry s.

As d_{m^2} is defined as the dimension of the space U_m, it also corresponds to the dimension of the space

$$vect\left\{C(\theta),\theta\in\Theta_m\right\}+vect\left\{C(\theta)^2,\theta\in\Theta_m\right\} , \qquad (15)$$