HAFNIANS, PERFECT MATCHINGS AND GAUSSIAN MATRICES

MARK RUDELSON, ALEX SAMORODNITSKY, AND OFER ZEITOUNI

ABSTRACT. We analyze the behavior of the Barvinok estimator of the hafnian of even dimension, symmetric matrices with non negative entries. We introduce a condition under which the Barvinok estimator achieves sub-exponential errors, and show that this condition is almost optimal. Using that hafnians count the number of perfect matchings in graphs, we conclude that Barvinok's estimator gives a polynomial-time algorithm for the approximate (up to subexponential errors) evaluation of the number of perfect matchings.

1. INTRODUCTION

The number of perfect matchings in a bipartite graph is given by the permanent of the bipartite adjacency matrix of the graph. Since computing the permanent is generally computationally hard [20], various algorithms have been proposed to compute it approximately. We mention in particular the MCMC algorithm of Jerrum-Sinclair-Vigoda [12], the Linial-Samorodnitsky-Wigderson rescaling algorithm [14] (denoted LSW in the sequel), and the Barvinok–Godsil-Gutman algorithm [9, 2];

the analysis of the latter algorithm was the subject of the previous work [18].

A more general (and hence hard) combinatorial problem is that of computing the number of perfect matchings in a graph with an even number of vertices. Let A denote the adjacency matrix of such a graph with $n = 2m$ vertices. The relevant combinatorial notion here is the *hafnian* [15], defined as

$$\text{haf}(A) = \frac{1}{m!2^m} \sum_{\sigma \in \mathcal{S}_n} \prod_{j=1}^{m} A_{\sigma(2j-1),\sigma(2j)}\,,$$

where $\mathcal{S}_n$ denotes the symmetric group on $[n]$. It is immediate to check, see e.g. [2], that

(1.1)   #perfect matchings in $A = \text{haf}(A)$.

Thus the interest in an efficient computation of $\text{haf}(A)$. As for the permanent, the exact computation of $\text{haf}(A)$ is computationally expensive. This problem of estimating the hafnian seems to be harder to attack than the corresponding problem for the permanent since many algorithms known for permanent approximation break down when extended to hafnians. In particular, the LSW rescaling algorithm [14] transforms the adjacency matrix of a graph to an almost doubly stochastic one. Yet, a non-trivial lower estimate of the hafnian of a doubly stochastic matrix is impossible, see [3]. Also, in contrast with the computation of the permanent, [12] points out that the proof of convergence of the MCMC algorithm breaks down for the approximate computation of the hafnian (unless the minimal degree is at least $n/2$, see [11]).

We consider in this paper the computation of $\text{haf}(A)$ for symmetric matrices with non-negative entries. Note that the diagonal entries play no role in the computation of $\text{haf}(A)$, and therefore in the rest of this paper we always assume that $A_{ii} = 0$ for all i.

In his seminal paper [2] discussing the Godsil-Gutman estimator for the permanent, Barvinok also introduces a probabilistic estimator of $\text{haf}(A)$ for a symmetric matrix A possessing non-negative entries. Let W be a real skew symmetric matrix with independent centered normal entries W_{ij} above the diagonal satisfying $\mathbb{E}W_{ij}^2 = A_{ij}$. In other words, let G^{skew} denote a skew symmetric matrix with independent $N(0,1)$ entries above the main diagonal. Let $W = W(A)$ denote the skew symmetric matrix with

(1.2) $W_{ij} = G_{ij}\sqrt{A_{ij}}, \ i < j$

and write $W = \mathcal{A} \odot G^{skew}$, where $\mathcal{A}$ denotes the element-wise square-root of A, i.e. $\mathcal{A}_{ij} = \sqrt{A_{ij}}$. Then,

$$(1.3) \qquad\qquad \mathrm{haf}(A) = E \det(W).$$

Thus, $\det(W)$, which is an easily computable quantity, is a consistent estimator for $\mathrm{haf}(A)$, and Barvinok [2] proceeds to prove that for any matrix A, $e^{-\gamma n}\mathrm{haf}(A) \leq \det(W) \leq C \cdot \mathrm{haf}(A)$ with high probability, where γ is Euler's constant. Other approaches to computing the hafnian include [3] (which however does not apply to adjacency matrices of nontrivial graphs), [5], where a deterministic algorithm of subexponential complexity is constructed and analyzed, and [4], where a random algorithm is analyzed but the precision of the algorithm depends in a complicated way on the number of perfect matchings.

Our goal in this paper is to analyze the performance of the Barvinok estimator for the hafnian. As in [18], establishing the concentration of a random determinant hinges on bounding the singular values of the Gaussian matrix W. This crucial step, however, essentially differs from [18] as W is skew-symmetric, and thus has less independence than the unrestricted matrix in [18]. Handling these dependences required different arguments for the smallest and the intermediate singular values. In the first case, we employ a conditioning argument tailored to take into account the structure of the graph (Lemmas 2.5, 2.6). The fact that the entries of W are real and thus the $(n-1) \times (n-1)$ main minors of it are degenerate plays a central role here. On the other hand, instead of developing a estimate for intermediate singular values as in [18] (a difficult task here due to the skew-symmetry), we use the fact that the imaginary-valued matrix iW is Hermitian, which allows us to use estimates from recent work [7] on the local semi-circle law.

To formulate our results, we introduce a notion of strong expansion for graphs. This notion strengthens the standard notion of vertex expansion assuming that sets having many connected components expand faster. For a set $J \subset [n]$ of vertices, denote by $\mathrm{Con}(J)$ the set of the connected components of J, and by $\partial(J)$ the boundary of J, that is $\partial(J) = \{i \in [n] \setminus J : \exists j \in J, (i, j) \text{ is an edge}\}$.

Definition 1.1. Let $\kappa \in (0, 1)$, and let $1 < m < n$. We say that the graph Γ is strongly expanding with parameter κ up to level m if for any set $J \subset [n]$ of vertices with $|J| \leq m$,

$$|\partial(J)| - |\mathrm{Con}(J)| \geq \kappa \cdot |J|.$$

In this definition and below, we use the following notational convention. Important parameters which appear in definitions and theorems

are denoted by Greek letters. Unimportant constants whose value may change from line to line are denoted c, c', C etc.

The simplest form of our results is in case A is the adjacency matrix of a d-regular graph.

Theorem 1.2. *Fix $\alpha, \kappa > 0$. Let A be the adjacency matrix of a d-regular graph Γ with $d \geq \alpha n + 2$. Assume that*

$$(1.4) \qquad \Gamma \text{ is } \kappa \text{ strongly expanding up to level } n(1-\alpha)/(1+\kappa/4).$$

    (1) *Then for any $\epsilon < 1/5$ and $D > 4$,*

$$(1.5) \qquad \mathbb{P}\left(|\log \operatorname{haf}(A) - \log \det(W)| > Cn^{1-\epsilon}\right) \leq n^{-D},$$

        where $C = C(\alpha, \kappa, D, \epsilon) > 0$.
    (2) *Fix $\delta > 0$. If, in addition to the above assumptions, the matrix A/d possesses a spectral gap δ, then for any $D > 4$,*

$$(1.6) \qquad \mathbb{P}\left(|\log \operatorname{haf}(A) - \log \det(W)| > Cn^{1/2}\log^{1/2} n\right) \leq n^{-D},$$

        where $C = C(\alpha, \kappa, D, \delta) > 0$.

(The assumption in (2) means that the modulus of the eigenvalues of A/d is either 1 or smaller than $1 - \delta$.) Theorem 1.2 is an immediate consequence of our more general Theorem 1.10 below.

We discuss the definition of strongly expanding graphs in Remark 1.6 below.

The extension of Theorem 1.2 to irregular graphs requires the notion of doubly stochastic scaling of matrices. We also need the notion of spectral gap for stochastic matrices.

Definition 1.3. A matrix A with non negative entries is said to possess a doubly stochastic scaling if there exist two diagonal matrices D_1, D_2 with positive entries such that the matrix $B = D_1 A D_2$ is doubly stochastic, that is $\sum_j B_{ij} = \sum_k B_{k\ell} = 1$ for all i, ℓ. We call such B a doubly stochastic scaling of A.

Definition 1.4. A symmetric stochastic matrix A is said to possess a *spectral gap* δ if there there do not exist eigenvalues of A in $(-1, -1 + \delta) \cup (1 - \delta, 1)$.

We will show below, see Corollary 3.3, that the adjacency matrix of a strongly expanding graph with appropriate lower bound on its minimal degree possesses a unique doubly stochastic scaling, with $D_1 = D_2$. We use this fact in the following theorem, where the required lower bound on the minimal degree is satisfied.

Theorem 1.5. *Fix $\alpha, \kappa, \vartheta > 0$. Let A be the adjacency matrix of a graph Γ whose minimal degree satisfies $d \geq \alpha n + 2$. Assume that*

(1.7)   *Γ is κ strongly expanding up to level $n(1-\alpha)/(1+\kappa/4)$,*

and

(1.8)

The doubly stochastic scaling B of A satisfies $\max_{i,j} B_{ij} \leq n^{-\vartheta}$.

   (1) *Then for any $\epsilon < 1/5$ and $D > 4$,*

(1.9) $$\mathbb{P}\left(|\log \operatorname{haf}(A) - \log \det(W)| > Cn^{1-\epsilon\vartheta}\right) \leq n^{-D}$$

   with $C = C(\alpha, \kappa, D, \epsilon) > 0$.
   (2) *Fix $\delta > 0$. If, in addition to the above assumptions, B possesses a spectral gap δ, then for any $D > 4$,*

(1.10) $$\mathbb{P}\left(|\log \operatorname{haf}(A) - \log \det(W)| > Cn^{1-\vartheta/2}\log^{1/2} n\right) \leq n^{-D},$$

   where $C = C(\alpha, \kappa, D, \delta, \vartheta) > 0$.

Condition (1.8) can be readily checked in polynomial time by applying the LSW scaling algorithm, stopped when its error is bounded above by n^{-1}. Indeed, at such time, the LSW algorithm output is a matrix $C = C(A)$ which is almost doubly stochastic in the sense that, with $B = B(A)$ denoting the doubly stochastic scaling of A, one has $\max_{ij} |B(A) - C(A)| < n^{-1}$. Because the maximal entry of B is at least n^{-1}, this implies that the maximal entries of B and of C are of the same order. Note also that the spectral gap condition in point (2) of Theorem 1.5, which depends only on the eigenvalues of B, can also be checked in polynomial time.

We note that for a given $0 < \vartheta < 1$, there exist stronger expansion conditions on the graph Γ which ensure that the maximal element in the doubly stochastic scaling of its adjacency matrix is of size at most $n^{-\vartheta}$. That is, if Γ satisfies these stronger properties, condition (1.8) is automatically satisfied. We refer to Section 6, Proposition 6.2, for details.

Conditions (1.7) and (1.8) play different roles in the proof. The first one is needed to establish the lower bound on the smallest singular value of W, and the second one guarantees that most of the singular values are greater than $n^{-\epsilon}$.

Remark 1.6. The definition of strongly expanding graphs (Definition 1.1 above) reminds one of that of a vertex expander. Yet, it is stronger in two senses. First, the strong expansion property takes into account the geometry of the set, requiring more rapid expansion for more "spread out" sets. Secondly, we want this expansion property

to hold for all sets of size relatively close to n, while for the classical expanders, the corresponding property is required only for sets with at most $n/2$ vertices. This may look unnatural at first glance. However, one may construct an example of a graph that can have the strong expansion property up to a level arbitrary close to 1, and yet the matrix corresponding to it may be degenerate with probability 1. Further, in Proposition 1.7 below, we construct a graph whose adjacency matrix barely misses the condition in Definition 1.1 and yet $\det(W)/\mathrm{haf}(A) \leq e^{-cn}$ with high probability for an appropriate $c > 0$.

Proposition 1.7. *Let $\delta > 0$. For any $N \in \mathbb{N}$, there exists a graph Γ with $M > N$ vertices such that*

$$(1.11) \quad \forall J \subset [M] \quad |J| \leq M/2 \;\Rightarrow\; |\partial(J)| - (1 - \delta)|Con(J)| \geq \kappa|J|$$

and

$$\mathbb{P}\left(\frac{\det(W)}{\mathbb{E}\det(W)} \leq e^{-cM} \right) \geq 1 - e^{-c'M}.$$

Here c, c', κ are constants depending on δ.

Our theorems on adjacency matrices are based on a general result pertaining to doubly stochastic symmetric matrices B with non-negative entries. We will consider matrices which have many relatively large entries. To formulate this requirement precisely, we introduce the notion of large entries graph.

Definition 1.8. Let A be a symmetric matrix with non-negative entries. For a parameter $\theta > 0$, define the *large entries graph* $\Gamma_A(\theta)$ by connecting the vertices $i, j \in [n]$ whenever $A_{ij} > \theta$. If A is the matrix of variances of entries of a skew symmetric matrix W, we will also refer to $\Gamma_A(\theta)$ as the *large variances graph* of W.

We will now formulate two theorems on the concentration of the hafnian of a skew symmetric matrix whose large variances graph satisfies a strong expansion condition.

Theorem 1.9. *Fix $\beta, \alpha, \vartheta, \kappa > 0$. Let B be a symmetric stochastic matrix of even size n with non negative entries, let $W = W(B)$ be as (1.2) (with B replacing A), and let $\Gamma = \Gamma_B(n^{-\beta})$ denote its large variances graph. Assume that:*

   (1) *The minimal degree of a vertex of Γ is at least $\alpha n + 2$.*
   (2) *Γ is κ strongly-expanding up to level $n(1 - \alpha)/(1 + \kappa/4)$.*
   (3) *$\max B_{ij} \leq n^{-\vartheta}$.*

Then, for any $\epsilon < 1/5$ and $D > 4$ there exists $C = C(\beta, \alpha, \kappa, \vartheta, \epsilon, D) > 0$ so that

$$(1.12) \qquad \mathbb{P}\left(|\log \mathrm{haf}(B) - \log \det(W)| > Cn^{1-\epsilon\vartheta}\right) \leq n^{-D} .$$

Somewhat tighter bounds are available if the matrix B possesses a spectral gap.

Theorem 1.10. *Assume the conditions of Theorem 1.9 and in addition assume that the matrix B has a spectral gap δ. Then, for any $D > 4$,*

$$(1.13) \qquad \mathbb{P}\left(|\log \mathrm{haf}(B) - \log \det(W)| > Cn^{1-\vartheta/2} \log^{1/2} n\right) \leq n^{-D}$$

The constant C here depends on all relevant parameters $\beta, \alpha, \kappa, \vartheta, \delta,$ and D.

The structure of the paper is as follows. In Section 2, we consider unit vectors that are close to vectors with small support and derive uniform small ball probability estimates for their images under the action of W. These estimates are used in Section 3 to obtain a lower bound for the smallest singular values of W. In Section 4, we provide local estimates for the empirical measure of eigenvalues of W. Section 5 is devoted to the proof of Theorems 1.9 and 1.10. Section 6 is devoted to the proof of a combinatorial lemma concerning the doubly stochastic scaling of adjacency matrices of strongly expanding graphs, which then is used in the proof of Theorem 1.5; in the section we also present sufficient conditions that ensure that (1.8) holds. Finally, in Section 7 we present the construction of the graph discussed in Proposition 1.7, and provide the proof of the latter.

Acknowledgment We thank Alexander Barvinok for many helpful discussions.

2. Compressible vectors

To establish the concentration for the determinant of the matrix W, we have to bound its smallest singular value. As is usual in this context, we view the smallest singular value of a matrix as the minimum of the norms of the images of unit vectors:

$$s_n(W) = \min_{x \in S^{n-1}} \|Wx\|_2 .$$

Before bounding the minimal norm over the whole sphere, let us consider the behavior of $\|Wx\|_2$ for a fixed $x \in S^{n-1}$. We begin with a small ball probability estimate, which is valid for any unit vector.

Lemma 2.1. *Let W be a skew-symmetric $n \times n$ matrix with independent, up to the symmetry restriction, normal entries. Assume that for*

any $j \in [n]$, there exist at least d numbers $i \in n$ such that $\mathrm{Var}(w_{ij}) \geq n^{-c}$. Then for any $x \in S^{m-1}$, and for any $t > 0$

$$\mathbb{P}\left(\|Wx\|_2 \leq tn^{-c'}\right) \leq (Ct)^{-d},$$

where C, c' depend on c only.

Proof. Let $x \in S^{n-1}$. Choose a coordinate $j \in [n]$ such that $|x_j| \geq n^{-1/2}$ and set $I = \{i \in [n] \mid \mathrm{Var}(w_{ij}) \geq n^{-c}\}$. Condition on all entries of the matrix W, except those in the j-th row and column. After this conditioning, for any $i \in I$ the i-th coordinate of the vector Wx is a normal random variable with variance $\mathrm{Var}(w_{i,j})x_j^2 \geq n^{-2c-1}$. Since the coordinates of this vector are conditionally independent, an elementary estimate of the Gaussian density yields for any $t > 0$

$$\mathbb{P}\left(\|Wx\|_2 < tn^{-c-1/2} \mid w_{ab}, \ a,b \in [n] \setminus \{j\}\right) \leq (Ct)^{|I|}.$$

By the assumption of the lemma, $|I| \geq d$. Integration with respect to the other variables completes the proof. $\qquad\square$

The next lemma is a rough estimate of the norm of a random matrix.

Lemma 2.2. *Let W be a be a skew-symmetric $n \times n$ matrix with independent, up to the symmetry restriction, normal entries. Assume that for any $i, j \in [n]$, $\mathrm{Var}(w_{ij}) \leq 1$. Then*

$$\mathbb{P}\left(\|W\| \geq n\right) \leq e^{-n}.$$

Lemma 2.2 follows from the estimate $\|W\|^2 \leq \|W\|_{HS}^2 = \sum_{ij=1}^n W_{ij}^2$, where the right side is the sum of squares of independent centered normal variables whose variances are uniformly bounded.

Of course, the estimate in Lemma 2.2 is very rough, but we can disregard a constant power of n in this argument. Lemma 2.2 allows us to extend the lower bound on the small ball probability from a single vector to a neighborhood of a small-dimensional subspace. To formulate it precisely, recall the definition of compressible and incompressible vectors from [16, 17].

Definition 2.3. For $m < n$ and $v < 1$, denote

$$\mathrm{Sparse}(m) = \{x \in S^{n-1} \mid |\mathrm{supp}(x)| \leq m\},$$

and

$$\mathrm{Comp}(m, v) = \{x \in S^{n-1} \mid \exists y \in \mathrm{Sparse}(m) \ \|x - y\|_2 \leq v\};$$
$$\mathrm{Incomp}(m, v) = S^{n-1} \setminus \mathrm{Comp}(m, v).$$

The next lemma uses a standard net argument to derive the uniform estimate for highly compressible vectors.

Lemma 2.4. *Let A be an $n \times n$ matrix satisfying the conditions of Lemmas 2.1 and 2.2. Then*

$$\mathbb{P}\left(\exists x \in \mathrm{Comp}\left(d/2, n^{-c}\right) \ \|Wx\|_2 \leq n^{-\bar{c}} \ and \ \|W\| \leq n\right) \leq e^{-d/2},$$

where $\bar{c}$ depends on c only.

Proof. Let $t > 0$ be a number to be chosen later, and set

$$\varepsilon = tn^{-c'-2},$$

where c' is the constant from Lemma 2.1. Then there exists an ε-net $\mathcal{N} \subset \mathrm{Sparse}(d/2)$ of cardinality

$$|\mathcal{N}| \leq \binom{n}{d/2} \cdot (3/\varepsilon)^{d/2} \leq \left(\frac{Cn}{tn^{-c'-2}}\right)^{d/2}.$$

By Lemma 2.1 and the union bound,

$$\mathbb{P}\left(\exists y \in \mathcal{N} \ \|Wy\|_2 \leq tn^{-c'}\right) \leq \left(\frac{Cn}{tn^{-c'-2}}\right)^{d/2} \cdot (Ct)^d \leq e^{-d/2},$$

provided that $t = n^{-c''}$ for an appropriately chosen $c'' > 0$.

Assume that for any $y \in \mathcal{N}$, $\|Wy\|_2 \geq tn^{-c'}$. Let $x \in \mathrm{Comp}(\delta/2, \varepsilon) = \mathrm{Comp}(\delta/2, n^{-c})$, and choose $y \in \mathcal{N}$ be such that $\|x - y\|_2 < 2\varepsilon$. If $\|W\| \leq n$, then

$$\|Wx\|_2 \geq \|Wy\|_2 - \|W\| \cdot \|x - y\|_2 \geq tn^{-c'} - n \cdot 2tn^{-c'-2} \geq n^{-\bar{c}}.$$

$\square$

Our next goal is to show that the small ball probability estimate propagates from strongly compressible vectors to moderately compressible ones. At this step, the assumption that the large variances graph is strongly expanding plays a crucial role. The strong expansion condition guarantees that the matrix W has enough independent entries to derive the small ball estimate for a single vector, despite the dependencies introduced by the skew-symmetric structure. The next simple lemma is instrumental in exploiting the independence that is still present.

Lemma 2.5. *Let $T = (V, E)$ be a finite tree with the root $r \in V$. Assume that to any $e \in E$ there corresponds a random variable X_e, and these variables are independent. Assume also that to any $v \in V$ there corresponds an event Ω_v, which depends only on those X_e for which $v \in e$. Suppose that for any $v \in V$ and any e_0 connected to v,*

$$\mathbb{P}\left(\Omega_v \mid \{X_e\}_{e \neq e_0}\right) \leq p_v$$

for some numbers $p_v \leq 1$. Then

$$\mathbb{P}\Big(\bigcap_{v \in V \setminus \{r\}} \Omega_v\Big) \leq \prod_{v \in V \setminus \{r\}} p_v.$$

Proof. We prove this lemma by induction on the depth of the tree. Assume first, that the tree has depth 2. Then the statement of the lemma follows from the fact that the events Ω_v, $v \in V \setminus \{r\}$ are independent.

Assume now that the statement holds for all trees of depths smaller than $k > 2$ and let T be a tree of depth k. Let V_r and E_r be the sets of all vertices and edges connected to the root of the tree. Then the events $\Omega_v, v \in V \setminus \{r\}$ conditioned on $x_e, e \notin E_r$ are independent. Therefore,

$$\mathbb{P}\Big(\bigcap_{v \in V \setminus \{r\}} \Omega_v\Big) = \mathbb{E}\,\mathbb{P}\Big[\bigcap_{v \in V \setminus \{r\}} \Omega_v \mid \{X_e\}_{e \notin E_r}\Big] \leq \prod_{v \in V_r} p_v \cdot \mathbb{P}\Big(\bigcap_{v \in V \setminus (V_r \cup \{r\})} \Omega_v\Big).$$

Note that the vertices $v \in V \setminus \{r\}$ form a forest with roots $v \in V_r$. Since the events $\cap_{v \in T_l} \Omega_v$ are independent for different trees T_l in the forest, the statement of the lemma follows by applying the induction hypothesis to each tree. $\square$

Using Lemma 2.5 and the strong expansion property of the large variances graph, we establish the small ball probability bound for the image of an incompressible vector.

Lemma 2.6. *Let $c > 0$. Let W be an $n \times n$ skew-symmetric centered Gaussian matrix. Assume that its large variances graph $\Gamma_W(n^{-c})$ satisfies the strong expansion condition with parameter $\kappa > 0$ up to level $m < n$. Let $t > 0, k \leq m$, and $v > 0$. Then for any $x \in Incomp(k, v)$,*

$$\mathbb{P}\left(\|Wx\|_2 \leq n^{-C}v \cdot t\right) \leq t^{(1+\kappa)k},$$

where C depends on c only.

Proof. For $i \in [n]$, define the event Ω_i by

$$\Omega_i = \{W : \ |(Wx)_i| \leq tn^{-(c+1)/2}vt\}.$$

Let $J(x) = \{j : \ |x_j| \geq n^{-1/2}v\}$. Since $x \in Incomp(k, v)$, $|J(x)| \geq k$. Indeed, let $y \in \mathbb{R}^n$ be the vector containing k largest in absolute value coordinates of x. If $|J(x)| \leq k$, then

$$\mathrm{dist}(x, \mathrm{Sparse}(k)) \leq \|x - y\|_2 \leq \left(\sum_{j \notin J(x)} x_j^2\right)^{1/2} \leq v.$$

Choose a subset $J \subset J(x)$ with $|J| = k$. For $i \in [n]$ set $p_i = t$ whenever $i \sim j$ for some $j \in J$; otherwise set $p_i = 1$. Then for any $j_0 \in J$ and for any $i_0 \sim j_0$,

$$\mathbb{P}\left(\Omega_{i_0} \mid W_{ij}, \ (i,j) \neq (i_0, j_0)\right) \leq t.$$

Indeed, $(Wx)_i$ is a normal random variable with variance at least

$$\mathrm{Var}(w_{i_0 j_0}) \cdot x_{j_0}^2 \geq n^{-c} \cdot n^{-1} v^2,$$

so the previous inequality follows from the bound on the maximal density.

To prove Lemma 2.6, we will use Lemma 2.5. To this end, we will construct a forest consisting of $L = |\mathrm{Con}(J)|$ trees with $|J| + |\partial(J)|$ vertices. Assume that such a forest is already constructed. The events $\cap_{i \in T_l} \Omega_i$ are independent for different trees T_l, $l = 1, \ldots, L$ in the forest. Hence,

$$\mathbb{P}\left(\|Wx\|_2 \leq tn^{-(c+1)/2} vt\right) \leq \mathbb{P}\left(|(Wx)_i| \leq tn^{-(c+1)/2} vt \text{ for all } i \in [n]\right)$$

$$\leq \prod_{l=1}^{L} \mathbb{P}\left(\bigcap_{i \in T_l} \Omega_i\right) \leq \prod_{l=1}^{L} t^{|T_l|-1} = t^{|J|+|\partial(J)|-L},$$

where we used Lemma 2.5 in the last inequality. Since by the strong expansion condition, $|J| + |\partial(J)| - L \geq (1+\kappa)|J|$, the last quantity is less than or equal to $t^{(1+\kappa)k}$ as required.

We proceed with the construction of the forest. At the first step, we construct a spanning tree $\tilde{T}_l$ for each connected component of the set J. These trees are, obviously, disjoint, and $\sum_{l=1}^{L} |\tilde{T}_l| = |J|$. Now, we have to add the vertices from $\partial(J)$ as leaves to these trees. We do this by induction on $j \in J$.

(1) Let $j \in J$ be the smallest number. Add all vertices $i \in \partial(J)$ connected to j to the tree containing j as the descendants of j.

(2) Let $j \in J$ be the smallest number, which has not been used in this process. Add all vertices $i \in \partial(J)$ connected to j, which have not been already added, to the tree containing j as its descendants.

Since any vertex in $\partial(J)$ is connected to some vertex in J, the whole set $\partial(J)$ will be added at the end of this process. Denote the trees obtained in this way by $T_1, \ldots, T_L$. The construction guarantees that these trees are disjoint. This finishes the construction of the forest and the proof of the lemma. $\qquad \square$

Similarly to Lemma 2.4, we extend the small ball probability result of Lemma 2.6 to a uniform bound using a net argument.

Lemma 2.7. *Let W be an $n \times n$ skew-symmetric Gaussian matrix. Assume that its large variances graph $\Gamma_W(n^{-c})$ satisfies the strong expansion condition with parameter $\kappa \in (0,1)$ up to level $m < n$. Then there exists a constant $C' > 0$ depending only on c and κ such that for any $t > 0, k \leq m$ and $v \in (0,1)$,*

$$\mathbb{P}\left(\exists x \in Incomp(k,v) \cap Comp((1+\kappa/2)k, (n^{-C'}v)^{8/\kappa})\right.$$
$$\left. \|Wx\|_2 \leq n \cdot (n^{-C'}v)^{8/\kappa} \text{ and } \|W\| \leq n\right) \leq e^{-k}.$$

Proof. The proof repeats that of Lemma 2.4, so we only sketch it. For $t > 0$, set

$$\varepsilon = n^{-C-2}vt,$$

where C is the constant from Lemma 2.6. Choose an ε-net $\mathcal{N}$ in $\text{Sparse}((1+\kappa/2)k) \cap \text{Incomp}(k,v)$ of cardinality

$$|\mathcal{N}| \leq \binom{n}{(1+\kappa/2)k} \cdot \left(\frac{3}{\varepsilon}\right)^{(1+\kappa/2)k} \leq \left(\frac{n^{\bar{c}}}{vt}\right)^{(1+\kappa/2)k},$$

where $\bar{c}$ depends only on c and κ. By the union bound,

$$\mathbb{P}\left(\exists x \in \mathcal{N} \ \|Wx\|_2 \leq n^{-C}vt \text{ and } \|W\| \leq n\right) \leq \left(\frac{n^{\bar{c}}}{vt}\right)^{(1+\kappa/2)k} \cdot t^{(1+\kappa)k}$$
$$\leq t^{(\kappa/4)k} \leq e^{-k},$$

provided that

$$t = \left(\frac{v}{en^{\bar{c}}}\right)^{4/\kappa}.$$

Using an appropriately defined $C' > 0$ depending only on c and κ, and approximation by the points of the ε-net, we derive from the previous inequality that

$$\mathbb{P}\left(\exists x \in \text{Incomp}(k,v) \cap \text{Sparse}((1+\kappa/2)k)\right.$$
$$\left. \|Wx\|_2 \leq 2n \cdot (n^{-C'}v)^{8/\kappa} \text{ and } \|W\| \leq n\right) \leq e^{-k}.$$

To complete the proof, notice that for any vector $y \in Incomp(k,v) \cap Comp((1+\kappa/2)k, (n^{-C'}v)^{8/\kappa})$, there is a vector $x \in Incomp(k,v) \cap \text{Sparse}((1+\kappa/2)k)$ such that $\|x - y\|_2 < (n^{-C'}v)^{8/\kappa}$. The Lemma now follows by again using approximation. $\square$

Lemmas 2.4 and 2.7 can be combined to treat all compressible vectors. In the statement, d_0 is a fixed, large enough universal positive integer.

Proposition 2.8. *Let W be an $n \times n$ skew-symmetric Gaussian matrix. Assume that its large variances graph $\Gamma_W(n^{-c})$ has minimal degree $d \geq d_0$ and satisfies the strong expansion condition with parameter $\kappa > 0$ up to the level $m < n$. Then there exists a constant $\phi(\kappa)$ depending on κ only so that, with $\rho = (n/d)^{\phi(\kappa)}$, one has*

$$\mathbb{P}\left(\exists x \in Comp((1 + \kappa/2)m, n^{-\rho}) \ \|Wx\|_2 \leq n^{-\rho+1}\right) \leq e^{-d/2}.$$

Remark 2.9. The proof below shows that it is enough to take

$$\phi(\kappa) = \frac{c}{\kappa} \log\left(\frac{C}{\kappa}\right).$$

Proof. Set $v_0 = n^{-c'}$, where $c' = \max(c, C')$, and c, C' are the constants from Lemmas 2.4 and 2.7. Let L be the smallest natural number such that

$$(d/2)(1 + \kappa/2)^L \geq m.$$

The definition of L implies

$$L \leq \frac{\log(m/d)}{\log(1 + \kappa/2)} \leq \frac{c}{\kappa} \log \frac{n}{d}.$$

For $l = 1, \ldots, L - 1$, define by induction $v_{l+1} = (n^{-C'}v_l)^{8/\kappa}$, where C' is the constant from Lemma 2.7. The definition of v_0 implies that $v_l \leq n^{-C'}$, so $v_{l+1} \geq v_l^{16/\kappa}$, and thus,

$$v_L \geq v_0^{(16/\kappa)^{L-1}} \geq n^{-\rho'} \quad \text{where } \rho' = \left(\frac{n}{d}\right)^{(c/\kappa)\cdot\log(C/\kappa)}.$$

We have

$$\mathrm{Comp}(m, v_L) \subset \mathrm{Comp}((d/2)(1 + \kappa/2)^L, v_L)$$
$$\subset \mathrm{Comp}(d/2, v_0) \cup$$
$$\bigcup_{l=1}^{L} \mathrm{Comp}((d/2)(1 + \kappa/2)^l, v_l) \setminus \mathrm{Comp}((d/2)(1 + \kappa/2)^{l-1}, v_{l-1}).$$

Lemmas 2.4 and 2.7 combined with the union bound imply

$$\mathbb{P}\left(\exists x \in \mathrm{Comp}(m, n^{-\rho'})\right.$$
$$\left.\|Wx\|_2 \leq n^{-\rho'+1} \text{ and } \|W\| \leq n\right) \leq e^{-d/2}.$$

Applying Lemma 2.7 once more, we derive the estimate

$$\mathbb{P}\left(\exists x \in \mathrm{Comp}((1 + \kappa/2)m, n^{-\rho})\right.$$
$$\left.\|Wx\|_2 \leq n^{-\rho+1} \text{ and } \|W\| \leq n\right) \leq e^{-d/2}$$

with $\rho = (\rho')^{16/\kappa}$.

The proposition follows from the previous inequality and Lemma 2.2. □

3. The smallest singular value

The main result of this section is the following lower bound for the smallest singular value of a Gaussian skew-symmetric matrix with a strongly expanding large variances graph.

Theorem 3.1. *Let $n \in \mathbb{N}$ be an even number. Let V be an $n \times n$ skew-symmetric matrix, and denote by Γ its large variances graph $\Gamma_V(n^{-c})$. Assume that*

(1) *$Var(v_{i,j}) \leq 1$ for all $i, j \in [n]$;*
(2) *the minimal degree of a vertex of Γ is at least $d \geq 2 \log n$;*
(3) *Γ is κ-strongly expanding up to level $\frac{n-d+2}{1+\kappa/4}$.*

Then
$$\mathbb{P}\left(s_n(V) \leq tn^{-\tau}\right) \leq nt + e^{-cd},$$
where $\tau = (n/d)^{\psi(\kappa)}$ for some positive $\psi(\kappa)$.

Remark 3.2. Tracing the proof of Theorem 3.1 and using Remark 2.9, one can show that it is enough to take
$$\psi(\kappa) = \frac{C'}{\kappa} \log\left(\frac{C}{\kappa}\right).$$

Proof of Theorem 3.1. To prove the theorem, we use the negative second moment identity. Let A be an $n \times n$ matrix with columns $A_1, \ldots, A_n$. For $j \in [n]$, let $h_j \in S^{n-1}$ be a vector orthogonal to all columns of A, except the j-th one. Then
$$\left\|A^{-1}\right\|_{HS}^2 = \sum_{j=1}^{n} (h_j^T A_j)^{-2}.$$

Hence,
$$s_n(A) = \frac{1}{\|A^{-1}\|} \geq \frac{1}{\|A^{-1}\|_{HS}} \geq n^{-1/2} \cdot \min_{j \in [n]} |h_j^T A_j|.$$

Let ρ be as in Proposition 2.8. The argument above shows that if we use the matrix V in place of A and define the unit vectors h_j, $j \in [n]$ as before, then the theorem would follow if the inequalities

(3.1) $$\mathbb{P}\left(|h_j^T V_j| \leq tn^{-\rho+c}\right) \leq t + e^{-cd}$$

hold for all $j \in [n]$. Indeed, the Theorem follows from (3.1) and the assumption on d by the union bound. We will establish inequality (3.1) for $j = 1$. The other cases are proved in the same way.

Let W be the $(n-1) \times (n-1)$ block of V consisting of rows and columns from 2 to n. The matrix W is skew-symmetric, and its large variances graph is the subgraph of Γ containing vertices $2, \ldots, n$. Therefore, Γ has properties (1), (2), (3) with slightly relaxed parameters. Indeed, property (1) remains unchanged. Property (2) is valid with d replaced by $d-1$. Property (3) is satisfied with parameter $\kappa/2$ in place of κ since for any $J \subset \Gamma \setminus \{1\}$, the boundary of J in Γ and in $\Gamma \setminus \{1\}$ differs by at most one vertex.

Recall that W is a skew-symmetric matrix of an odd size. This matrix is degenerate, so there exists $u \in S^{n-2}$ such that $Wu = 0$. This allows us to define the vector $h \in S^{n-1}$ orthogonal to the columns $V_2, \ldots, V_n$ of the matrix V by

$$h = \begin{pmatrix} 0 \\ u \end{pmatrix}.$$

Define the event Ω by

$$\Omega = \{W : \ \exists u \in \mathrm{Comp}(n-d+1, n^{-\rho}) \ Wu = 0\}.$$

The graph $\Gamma_W(n^{-c}) \setminus \{1\}$ is $(\kappa/4)$ strongly expanding up to level $\frac{n-d+1}{1+\kappa/4}$. By Proposition 2.8, $\mathbb{P}(\Omega) \leq e^{-d/2}$. Condition on the matrix $W \in \Omega^\complement$. After the conditioning, we may assume that $u \in \mathrm{Incomp}(n-d+1, n^{-\rho})$, and so the set $J = \{j : \ |u_j| \geq n^{-\rho-1/2}\}$ has at least $n-d+1$ elements. Since the degree of the vertex $\{1\}$ in the large variances graph of V is at least d, this means that there exists a $j \in J$ for which $\mathrm{Var}(v_{j1}) \geq n^{-c}$. Therefore, conditionally on W, $h^T V_1 = \sum_{j=2}^{n} u_j v_j$ is a normal random variable with variance

$$\mathrm{Var}(h^T V_1) = \sum_{j=2}^{n} u_j^2 \cdot \mathrm{Var}(v_j) \geq n^{-2\rho-1} \cdot n^{-c}.$$

The bound on the density of a normal random variable implies

$$\mathbb{P}\left(|h^T V_1| \leq C n^{-\rho-c/2-1/2} t \mid W \in \Omega^\complement\right) \leq t.$$

Finally,

$$\begin{aligned} &\mathbb{P}\left(|h^T V_1| \leq C n^{-\rho-c/2-1/2} t\right) \\ &\leq \mathbb{P}\left(|h^T V_1| \leq C n^{-\rho-c/2-1/2} t \mid W \in \Omega^\complement\right) + \mathbb{P}\left(W \in \Omega\right) \\ &\leq t + e^{-d/2}. \end{aligned}$$

This completes the proof of (3.1) for $j = 1$. Since the proof for the other values of j is the same, it proves Theorem 3.1. $\qquad \square$

An immediate corollary of Theorem 3.1 is the following.

Corollary 3.3. *Let A be the adjacency matrix of a graph Γ which satisfies*

(1) *the minimal degree of a vertex of Γ is at least $d > 2\log n$;*
(2) *Γ is κ-strongly expanding up to level $\frac{n-d+2}{1+\kappa/4}$.*

Then, A possesses a unique doubly stochastic scaling $B = DAD$ and the graph Γ possesses a perfect matching.

Proof of Corollary 3.3. We begin by showing that a perfect matching in Γ exists. Assume otherwise. Then $s_n(A \odot G) = 0$ since $\mathbb{E}\det(A \odot G) = \mathrm{haf}(A) = 0$. The latter equality contradicts Theorem 3.1.

To show that A possesses a doubly stochastic scaling, choose an edge $e = (u, v)$ in Γ and create a graph Γ' by erasing u, v, and all edges attached to them from Γ. The graph Γ' satisfies assumptions (1) and (2) in the statement, with slightly smaller constants κ, d. Thus, Γ' possesses a perfect matching. This implies that for any edge in Γ there exists a perfect matching containing that edge. By Bregman's theorem [6, Theorem 1], this implies that A possesses a unique doubly stochastic scaling $B = D_1 A D_2$. The fact that $D_1 = D_2$ follows from the strict convexity of relative entropy and the characterization of the doubly stochastic scaling as its minimizer, see [6, Eq. (7)]. $\square$

4. Local bound on eigenvalues density

In this section, we prove a general bound on the crowding of eigenvalues at 0 for a class of Hermitian matrices whose variance matrix is doubly-stochastic. The results are somewhat more general than our needs in the rest of the paper and may be of independent interest, and therefore we introduce new notation.

Let X denote an $n \times n$ matrix, Hermitian (in the sense that $X_{ij}^* = X_{ji}$), with entries $\{X_{ij}\}_{i \leq j}$ that are independent zero mean random variables. (In our application, the X_{ij} variables are all Gaussian.). Following [7], we set $s_{ij} = E|X_{ij}|^2$ and $\zeta_{ij} = X_{ij}/\sqrt{s_{ij}}$ (with $\zeta_{ij} = 0$ if $s_{ij} = 0$). We assume that the variables ζ_{ij} possess uniformly bounded p moments for all $p > 0$. Finally, we denote the eigenvalues of the matrix X by $\lambda_1(n) \geq \lambda_2(n) \geq \ldots \geq \lambda_n(n)$, and use $L_n = n^{-1}\sum_{i=1}^n \delta_{\lambda_i(n)}$ for the empirical measure of eigenvalues.

We assume that $\sum_j s_{ij} = 1$, and, to avoid trivialities, that the matrix $S = \{s_{ij}\}$ is irreducible (otherwise, the matrix X can be decomposed in blocks due to the symmetry).

Let $M = (\max_{ij} s_{ij})^{-1}$. We assume the following.

Assumption 4.1. For some $\vartheta \in (0, 1]$ one has that $M \geq n^\vartheta$.

With Assumption 4.1, we have the following proposition.

Proposition 4.2. *With notation and assumptions as in the setup above, fix $\epsilon < 1/5$. If Assumption 4.1 holds then for every $D > 0$ there exists $n_0 = n_0(\epsilon, D)$ such that for any $n > n_0$, and with*

$$\mathcal{N}(\eta) = |\{i : \lambda_i(n) \in (-\eta, \eta)\}|,$$

one has

(4.1) $$\mathbb{P}\left(\exists \eta \geq M^{-\epsilon} \ \mathcal{N}(\eta) > Cn \cdot \eta\right) \leq n^{-D} .$$

Proof. We will use [7, Theorem 2.3], a simplified form of which we quote below after introducing some notation. Following the notation in [7], we let $m(z)$ denote the Stieltjes transform of the semicircle law, $m(z) = (-z + \sqrt{z^2 - 4})/2$, and set

$$\Gamma(z) = \|(1 - m(z)^2 S)^{-1}\|_{\ell^\infty \to \ell^\infty}.$$

Note that with $z = i\eta$ one has by [7, (A.1)] that for some universal constant C,

(4.2) $$\Gamma(i\eta) \leq \frac{C \log n}{\eta} .$$

Introduce now, similarly to [7, (2.14)], for a parameter $\gamma > 0$,

$$\tilde{\eta} = \min\left\{u > 0 : \frac{1}{Mu} \leq \min\left\{\frac{M^{-\gamma}}{\Gamma(i\eta')^3}, \frac{M^{-2\gamma}}{\Gamma(i\eta')^4 \cdot \Im(m(i\eta'))}\right\}, \right.$$
$$\left. \text{for all } \eta' \in [u, 10]\right\}.$$

(Note that we do not use $\tilde{\Gamma}_N(z)$ as in [7] since we only need the relation $\Gamma \geq \tilde{\Gamma}_N$.) Note that $\Im(m(i\eta')) = (\sqrt{4 + \eta'^2} - \eta')/2$ is bounded above and below by a universal constant for $\eta' \in [0, 10]$. Hence, using (4.2), we get that

(4.3) $$\tilde{\eta} \leq C(M^{2\gamma - 1}(\log n)^4)^{1/5} := \bar{\eta},$$

for some universal constant C. For given $\varepsilon < 1/5$, we will chose $\gamma \in (0, 1/2)$ and $n_0 = n_0(\varepsilon, D)$ so that $\bar{\eta} \leq M^{-\varepsilon}$ whenever $n > n_0$.

Denote by $m_n(z) = n^{-1} \sum_{i=1}^n \frac{1}{\lambda_i - z}$ the Stieltjes transform of the empirical measure of eigenvalues of X. We have the following.

Theorem 4.3. [7, Theorem 2.3] *For any $\gamma \in (0, 1/2)$, any $\epsilon' > 0$ and any $D > 0$ there exists an $n_0 = n_0(D, \gamma, \epsilon')$ so that, uniformly in $\eta \in [\tilde{\eta}, 10]$, and for all $n > n_0$,*

(4.4) $$\mathbb{P}\left(|m_n(i\eta) - m(i\eta)| > \frac{n^{\epsilon'}}{M\eta}\right) \leq n^{-D} .$$

Fix $\eta \geq \bar{\eta}$. Let $\mathcal{A}$ denote the complement of the event in (4.4). Assume that $\mathcal{A}$ occurs. Using the uniform boundedness of $m(i\eta)$ and inequality (4.3), we obtain

$$\Im m_n(i\eta) \leq C + \frac{n^{\epsilon'}}{M\eta} \leq C + C\frac{n^{\epsilon'}}{M \cdot (M^{2\gamma-1}(\log n)^4)^{1/5}} \, .$$

Choosing γ and ϵ' small enough, we can guarantee that the right side in the last display is uniformly bounded in n. With such choice,

$$C \geq \int \frac{\eta}{\lambda^2 + \eta^2} dL_n(\lambda) \geq \frac{1}{2\eta} \int_{-\eta}^{\eta} dL_n(\lambda) = \frac{1}{2\eta} L_n([-\eta, \eta])$$

provided that $\mathcal{A}$ occurs. This means that

$$\mathbb{P}\left(\mathcal{N}(\eta) > Cn \cdot \eta\right) \leq n^{-D} \, .$$

To derive (4.1) from the previous inequality, one can use the union bound over $\eta = 1/k$ with $k \in \mathbb{N}$, $M^{\varepsilon} \geq k \geq 1$.   □

A better estimate can be obtained if one assumes a spectral gap. First, we have the following.

Lemma 4.4. *S has exactly one eigenvalue at $+1$ and at most one eigenvalue at -1.*

Proof. The claim concerning the eigenvalue at 1 is the Perron–Frobenius Theorem. To check the claim on the eigenvalues at -1, consider S^2. It may be reducible, but at most to 2 blocks. Indeed, suppose there are 3 disjoint blocks A_1, A_2, A_3, i.e. disjoint subsets A_i of $[N]$, $i = 1, 2, 3$, so that for all $a \in A_i, b \in A_j$ with $i \neq j$ one has $S_{a,b}^2 = 0$. By the irreducibility of S, there is a path of odd length connecting A_1 and A_2, and similarly there is a path of odd length connecting A_2 and A_3. Hence, there is a path of even length connecting A_1 and A_3, in contradiction with the block disjointness of S^2. The claim now follows by applying the Perron–Frobenius Theorem to each of the blocks of S^2.   □

By Lemma 4.4, if S has a spectral gap then the eigenvalues at 1 and -1 (if the later exists) are unique and isolated.

Proposition 4.5. *With notation and assumptions as in the setup above, fix $\epsilon < 1$. If Assumption 4.1 holds and S possesses a spectral gap δ then for every $D > 0$ there exists $n_0 = n_0(\epsilon, D, \delta)$ such that for any $n > n_0$, with*

$$\mathcal{N}(\eta) = |\{i : \lambda_i(n) \in (-\eta, \eta)\}|,$$

one has

(4.5) $$\mathbb{P}\left(\exists \eta \geq M^{-\epsilon} \;\; \mathcal{N}(\eta) > n \cdot \eta\right) \leq n^{-D} \, .$$

The proof is identical to that of Proposition 4.2, using [1, Theorem 1.1] instead of [7, Theorem 1.2]. We omit the details.

5. CONCENTRATION OF THE HAFNIAN OF A RANDOM MATRIX.

In this section we prove Theorems 1.9 and 1.10. Both results follow from the concentration of the Gaussian measure for Lipschitz functions. To this end, we consider a Gaussian vector $G = (G_{ij})_{1 \leq i < j \leq n} \in \mathbb{R}^{n(n-1)/2}$ and use it to form the skew-symmetric matrix G^{skew}. However, the function $F(G) = \log \det(\mathcal{B} \odot G^{skew})$ (where $\mathcal{B}_{ij} = \sqrt{B_{ij}}$) is not Lipschitz. To overcome this obstacle, we write

$$\log \det(\mathcal{B} \odot G^{skew}) = \sum_{j=1}^{n} \log s_j(\mathcal{B} \odot G^{skew})$$

and use Theorem 3.1 and Proposition 4.2 to obtain lower bounds on the singular values which are valid with probability close to 1. On this event, we replace the function log by its truncated version, which makes it Lipschitz with a controlled Lipschitz constant. Then an application of the Gaussian concentration inequality yields the concentration of the new truncated function about its expectation. This expectation is close to $\mathbb{E} \log \det(\mathcal{B} \odot G^{skew})$. Recall that instead of the concentration about this value, we want to establish the concentration about $\log \operatorname{haf}(B) = \log \mathbb{E} \det(\mathcal{B} \odot G^{skew})$. In other words, we have to swap the expectation and the logarithm and estimate the error incurred in this process. This will be achieved due to the fast decay of the tail in the concentration inequality.

Proof of Theorem 1.9. The proof proceeds as in [18, Section 7]. (The argument can be traced back to [8].) Without loss of generality, we may assume that $n > n_0$, where $n_0 = n_0(\varepsilon, D)$ appears in Proposition 4.2. Indeed, if $n \leq n_0$, we can choose the constant C in the formulation of the theorem appropriately large, so that $Cn_0^{1-\varepsilon\vartheta} \geq n_0$. In this case, Theorem 1.9 follows from Barvinok's theorem.

Fix $\epsilon < 1/5$, $D > 4$ as in the statement of the theorem, and $t = n^{-(D+1)}$. With $\tau = \tau(\kappa, \alpha)$ as in Theorem 3.1 (with X replaced by W) and $\mathcal{N}(\eta) = \mathcal{N}(\eta)(W)$ as in Proposition 4.2, introduce the events

$$\mathcal{W}_1 = \{s_n(W) \leq tn^{-\tau}\}, \quad \mathcal{W}_2 = \{\mathcal{N}(n^{-\varepsilon\vartheta}) \geq n^{1-\epsilon\vartheta}\}, \quad \mathcal{W} = \mathcal{W}_1 \cup \mathcal{W}_2.$$

By Theorem 3.1 and Proposition 4.2 we have that for all $n > n_0$,

$$(5.1) \qquad\qquad \mathbb{P}(\mathcal{W}) \leq 3n^{-D}.$$

Let $\widetilde{\det}(W) = \prod_i(|\lambda_i(W)| \vee n^{-\epsilon\vartheta})$. Note that on $\mathcal{W}^{\complement}$ we have that

$$(5.2) \qquad |\log\widetilde{\det}(W) - \log\det(W)| \leq C(D)n^{1-\epsilon\vartheta}\log n.$$

Set $U = \log\widetilde{\det}(W) - \mathbb{E}\log\widetilde{\det}(W)$. We next derive concentration results for U. The map $(\lambda_i(W))_{i=1}^N \to \log\widetilde{\det}(W)$ is Lipschitz with constant $n^{1/2+\epsilon\vartheta}$. Therefore, by standard concentration for the Gaussian distribution, see [10, 13], using that the variance of the entries of W is bounded above by $n^{-\vartheta}$, we have for some universal constant C and any $u > 0$,

$$(5.3) \qquad \mathbb{P}\left(|U| > u\right) = \mathbb{P}\left(|\log\widetilde{\det}(W) - \mathbb{E}\log\widetilde{\det}(W)| > u\right)$$

$$\leq \exp\left(-\frac{Cu^2}{n^{1+(2\epsilon-1)\vartheta}}\right).$$

Therefore,

$$(5.4) \quad \mathbb{E}(e^{|U|}) \leq 1 + \int_0^\infty \exp\left(u - \frac{Cu^2}{n^{1+(2\epsilon-1)\vartheta}}\right)\,du \leq \exp\left(n^{1+(2\epsilon-1)\vartheta}\right).$$

In particular, we obtain that

$$(5.5) \qquad \mathbb{E}\log\widetilde{\det}W \leq \log\mathbb{E}\widetilde{\det}W \leq \mathbb{E}\log\widetilde{\det}W + n^{1+(2\epsilon-1)\vartheta}.$$

The first inequality above follows from Jensen's inequality, and the second one from (5.4).

We can now complete the proof of the theorem. We have by Markov's inequality that

$$(5.6) \qquad \mathbb{P}\left(\log\det(W) - \log\mathbb{E}\det(W) > n^{1-\epsilon\vartheta}\log n\right) \leq e^{-n^{1-\epsilon\vartheta}\log n}.$$

On the other hand, note that $\mathbb{E}\det(W) \leq \mathbb{E}\widetilde{\det}(W)$. Therefore, with $C(D)$ as in (5.2),

$$\mathbb{P}\left(\log\det(W) - \log\mathbb{E}\det(W) \leq -(C(D)+2)n^{1-\epsilon\vartheta}\log n\right)$$
$$\leq \quad \mathbb{P}\left(\log\det(W) - \log\mathbb{E}\widetilde{\det}(W) \leq -(C(D)+2)n^{1-\epsilon\vartheta}\log n\right)$$
$$\leq \quad \mathbb{P}\left(\log\widetilde{\det}(W) - \log\mathbb{E}\widetilde{\det}(W) \leq -2n^{1-\epsilon\vartheta}\log n\right) + P(\mathcal{W}^{\complement}),$$

where (5.2) was used in the last display. Using now (5.1) and the upper bound in (5.5), we get

$$\mathbb{P}\left(\log\det(W) - \log\mathbb{E}\det(W) \leq -(C(D)+2)n^{1-\epsilon\vartheta}\log n\right)$$
$$\leq \quad 3n^{-D} + \mathbb{P}\left(\log\widetilde{\det}(W) - \mathbb{E}\log\widetilde{\det}(W) \leq -2n^{1-\epsilon\vartheta}\log n + n^{1+(2\epsilon-1)\vartheta}\right).$$

Using that $\epsilon < 1/5$ and applying (5.3) we conclude that

$$\mathbb{P}\left(\log\det(W) - \log\mathbb{E}\det(W) \leq -(C(D)+2)n^{1-\epsilon\vartheta}\log n\right) \leq 4n^{-D}.$$

Together with (5.6), it yields

$$\mathbb{P}\left(|\log\det(W) - \log\mathbb{E}\det(W)| \leq (C(D)+2)n^{1-\epsilon\vartheta}\log n\right) \leq 5n^{-D}.$$

To obtain the statement of the theorem, we prove the previous inequality with $\varepsilon' \in (\varepsilon, 1/5)$ instead of ε, and then choose $C' > 0$ such that $(C(D)+2)n^{1-\varepsilon\vartheta}\log n \leq C'n^{1-\varepsilon'\vartheta}$.   □

The proof of Theorem 1.10 is similar to that of 1.9. However, to exploit the tighter bounds on the intermediate singular values provided by Proposition 4.5, we use a different truncation, redefining the function $\log\widetilde{\det}(\cdot)$ and estimate its Lipschitz constant more accurately.

Proof of Theorem 1.10. Fix $\varepsilon \in (1/2, 1)$. As in the proof of Theorem 1.9, we may assume that $n > n_0(\varepsilon, D, \delta)$, where $n_0(\varepsilon, D, \delta)$ was introduced in Proposition 4.5. The inequality (4.5) can be rewritten as

$$(5.7) \qquad \mathbb{P}\left(\exists k \geq n^{1-\varepsilon\vartheta} \quad s_{n-k}(W) \leq c\frac{k}{n}\right) \leq n^{-D}.$$

This inequality can be used to bound the Lipschitz constant of the truncated logarithm. Let

$$(5.8) \qquad\qquad\qquad m_0 \geq n^{1-\varepsilon\vartheta}$$

be a number to be chosen later. Define

$$(5.9) \qquad \log\widehat{\det}(W) = \sum_{k=1}^{n} \phi_k(s_{n-k}(W)) = \sum_{k=1}^{n} \log(s_{n-k}(W) \vee \varepsilon_k),$$

where

$$\varepsilon_k = \begin{cases} c\frac{m_0}{n}, & \text{for } k < m_0 \\ c\frac{k}{n}, & \text{for } k \geq m_0 \end{cases}$$

Denote for a moment $N = n(n-1)/2$. For a vector $Y \in \mathbb{R}^N$ consider the $n \times n$ skew symmetric matrix Y^{skew} whose entries above the main diagonal equal to the corresponding entries of Y. Let $\mathcal{B}$ be the $n \times n$ matrix whose entries are square roots of the corresponding entries of B. Note that the function $F : \mathbb{R}^N \to \mathbb{R}$ defined by

$$F(G) = \log\widehat{\det}(\mathcal{B} \odot G^{skew})$$

is the composition of three functions: $F = F_3 \circ F_2 \circ F_1$, where
  (1) $F_1 : \mathbb{R}^N \to \mathbb{R}^{n^2}$, $F_1(G) = \mathcal{B} \odot G^{skew}$, whose Lipschitz constant does not exceed $n^{-\vartheta/2}$;
  (2) $F : \mathbb{R}^{n^2} \to \mathbb{R}^n_+$ defined by $F(W) = (s_1(W), \ldots, s_n(W))$, which is 1-Lipschitz;
  (3) $F_3 : \mathbb{R}^n_+ \to \mathbb{R}$, $F_3(x_1, \ldots, x_n) = \sum_{k=1}^{n} \phi_k(x_{n-k})$, where ϕ_k is defined in (5.9).

By the Cauchy-Schwarz inequality,

$$\|F_3\|_{Lip} \leq \left(\sum_{k=1}^{n} \|\phi_k\|_{Lip}^2\right)^{1/2} \leq \left(\left(\frac{n}{cm_0}\right)^2 \cdot m_0 + \sum_{k>m_0} \left(\frac{n}{ck}\right)^2\right)^{1/2}$$

$$\leq C\frac{n}{\sqrt{m_0}}.$$

Therefore,

$$\|F\|_{Lip} \leq C\frac{n^{1-\vartheta/2}}{\sqrt{m_0}}.$$

Applying the standard Gaussian concentration for Lipschitz functions, we obtain

$$(5.10) \quad \mathbb{P}\left(|\log\widehat{\det}(W) - \mathbb{E}\log\widehat{\det}(W)| \geq u\right) \leq 2\exp\left(-\frac{u^2}{2\|F\|_{Lip}^2}\right)$$

$$\leq 2\exp(-cm_0 n^{\vartheta-2}u^2)$$

which replaces formula (5.8) in the proof of Theorem 1.9. Arguing as in the proof of that theorem, we obtain

$$(5.11) \qquad \mathbb{E}\log\widehat{\det}(W) \leq \log\mathbb{E}\widehat{\det}(W) \leq \mathbb{E}\log\widehat{\det}(W) + C_0\frac{n^{2-\vartheta}}{m_0}$$

from the inequality above.

Let $C' = D + \tau(\kappa, \alpha)$, where $\tau(\kappa, \alpha)$ is as in Theorem 3.1. Set

$$\mathcal{W}_1 = \left\{s_n(W) \leq n^{-C'}\right\}, \quad \mathcal{W}_2 = \left\{\exists k \geq n^{1-\varepsilon\vartheta} \ \ s_{n-k}(W) \leq c\frac{k}{n}\right\},$$

and let $\mathcal{W} = \mathcal{W}_1 \cup \mathcal{W}_2$. Then Theorem 3.1 and (5.7) imply $\mathbb{P}(\mathcal{W}) \leq n^{-D}$. On $\mathcal{W}^{\complement}$ we have

$$(5.12) \qquad |\log\widehat{\det}(W) - \log\det(W)| \leq Cm_0\log n,$$

which plays the role of (5.7). Arguing as in the proof of Theorem 1.9, we show that

$$\mathbb{P}\left(\log\det(W) - \log\mathbb{E}\det(W) \leq -\left(Cm_0\log n + 2C_0\frac{n^{2-\vartheta}}{m_0}\right)\right)$$

$$\leq \mathbb{P}\left(\log\widehat{\det}(W) - \log\mathbb{E}\widehat{\det}(W) \leq -2C_0\frac{n^{2-\vartheta}}{m_0}\right) + \mathbb{P}(\mathcal{W})$$

$$\leq \mathbb{P}\left(\log\widehat{\det}(W) - \mathbb{E}\log\widehat{\det}(W) \leq -C_0\frac{n^{2-\vartheta}}{m_0}\right) + \mathbb{P}(\mathcal{W})$$

$$\leq \exp\left(-C'\frac{n^{2-\vartheta}}{m_0}\right) + n^{-D} \leq 2n^{-D}.$$

Here, the first inequality follows from (5.12) and $\log \mathbb{E} \det(W) \leq \log \widehat{\mathbb{E} \det}(W)$, the second one from the upper bound in (5.11), and the third one from (5.10).

We select the optimal $m_0 = Cn^{1-\vartheta/2} \log^{-1/2} n$ in the inequality above. Since $\varepsilon > 1/2$, condition (5.8) holds for sufficiently large n.

$$\mathbb{P}\left(\log \det(W) - \log \mathbb{E} \det(W) \leq -Cn^{1-\vartheta/2} \log^{1/2} n\right) \leq 2n^{-D}.$$

Combining this with the bound

$$\mathbb{P}\left(\log \det(W) - \log \mathbb{E} \det(W) \geq Cn^{1-\vartheta/2} \log^{1/2} n\right) \leq n^{-D}$$

following from Markov's inequality, we complete the proof. $\square$

6. Doubly stochastic scaling and proof of Theorem 1.5

To prove Theorem 1.5, we have to scale the adjacency matrix of the graph in order to apply Theorems 1.9 and 1.10. The existence of such scaling has been already established in Corollary 3.3. We will show now that the smallest non-zero entry of the scaled adjacency matrix is at least polynomial in n. This crucial step in the proof of Theorem 1.5 allows us to conclude that the large entries graph of the scaled matrix coincides with the original graph.

Proposition 6.1. *Fix $\alpha, \kappa > 0$. Let A be the adjacency matrix of a graph Γ whose minimal degree satisfies $d \geq \alpha n + 2$. Assume that*

$$(6.1) \qquad \Gamma \text{ is } \kappa \text{ strongly expanding up to level } n(1-\alpha)/(1+\kappa/4),$$

Then there exists a constant ν so that A possesses a doubly stochastic scaling $B = DAD$ with

$$\min_i D_{ii} \geq n^{-\nu}.$$

In particular, under the assumptions of Proposition 6.1, we have that

$$(6.2) \qquad\qquad B_{ij} \geq n^{-2\nu} \text{ whenever } B_{ij} > 0.$$

Before describing the proof of Proposition 6.1, let us state a complementary claim, which says that under stronger expansion conditions on Γ we can guarantee that the entries in its scaled adjacency matrix are *polynomially small*. This ensures that under this stronger expansion property, condition (1.8) in Theorem 1.5 is automatically satisfied.

In what follow, if X is a set of vertices in a graph then $E(X, X)$ denotes those edges in the graph connecting vertices in X.

Proposition 6.2. *Fix $\alpha, \kappa, \varepsilon > 0$. There exists a constant θ depending only on α and κ such that the following holds.*

Let A be the adjacency matrix of a graph Γ whose minimal degree satisfies $d \geq \alpha n + 2$. Assume that for any subset X of vertices satisfying $|X| \geq \frac{d}{4}$ and $|E(X, X)| \leq \theta \cdot n^{1+\varepsilon}$, it holds that

$$|\partial_s X| \geq (1 + \kappa) \cdot |X|$$

where $\partial_s X$ denotes the set of external neighbors of X such that any $y \in \partial_s X$ has at least $\frac{d|X|}{10n}$ neighbors in X.

Then A possesses a doubly stochastic scaling $B = DAD$ with

$$\max_i D_{ii} \leq n^{-\varepsilon/2} \, .$$

In particular, under the assumptions of Proposition 6.2, we have that

(6.3) $$B_{ij} \leq n^{-\varepsilon} \text{ whenever } B_{ij} > 0 \, .$$

To prove Proposition 6.1, we argue by contradiction. Assume that one of the diagonal entries, say D_{11}, is smaller than $n^{-\nu}$, where $\nu = \nu(\alpha, \kappa)$ will be chosen at the end of the proof. The double stochasticity of the scaled matrix implies that there exists a neighbor $i \sim 1$ for which the corresponding entry of the scaling matrix D is large. In fact, we can prove this for more than one entry. In Lemma 6.5 we construct a set $X = X_0$ of vertices of cardinality at least $d/2$ such that the corresponding entries of the scaling matrix are greater than $(1/2)n^{\nu-2}$. We use this as a base of induction. In Lemma 6.6, we show that there exists a set X_1 of vertices of cardinality $|X_1| \geq (1+\beta)|X_0|$ containing X_1 such that all entries of the scaling matrix corresponding to X_1 are still polynomially large. Proceeding by induction, we construct an increasing sequence of sets $X_0 \subset X_1 \subset \cdots \subset X_l$ such that $|X_l| \geq (1+\beta)^l|X_0|$, and all diagonal entries corresponding to the vertices of X_l are greater than 1. The number of induction steps l which we are able to perform will depend on ν. If ν is chosen large enough, then we will get $(1+\beta)^l|X_0| > n$, reaching the desired contradiction.

The proof of Proposition 6.2 is very similar. Assume, towards contradiction, that, say, D_{nn} is larger than $n^{-\varepsilon/2}$. By the double stochasticity of the scaled matrix, there exists a set $\mathcal{A}$ of neighbors $i \sim n$, for which the corresponding entries of the scaling matrix D are small. Using again the double stochasticity of the scaled matrix produces a set $X = X_0$ of vertices of cardinality at least $d/4$ such that the corresponding entries of the scaling matrix are greater than $(\alpha/8)n^{-\varepsilon/2}$. We use this as induction base. In Lemma 6.7, we show that there exists a set X_1 of vertices of cardinality $|X_1| \geq (1+\gamma)|X_0|$ containing X_1 such

that all entries of the scaling matrix corresponding to X_1 are still large. Proceeding by induction, we construct an increasing sequence of sets $X_0 \subset X_1 \subset \cdots \subset X_l$ such that $|X_l| \geq (1 + \gamma)^l |X_0|$, and all diagonal entries corresponding to the vertices of X_l are greater than $\Omega\left(n^{-\varepsilon/2}\right)$. The number of induction steps l which we are able to perform will depend on θ. If $\theta = \theta(\alpha, \kappa)$ is chosen large enough, then we will get $(1 + \gamma)^l |X_0| > n$, reaching contradiction.

Proof of Proposition 6.1. Without loss of generality, we assume throughout that the constants α and κ are small enough so that

$$(6.4) \qquad\qquad (1 - \alpha)/(1 - \kappa/4) > 1/2 \,.$$

By Corollary 3.3, A possesses a doubly stochastic scaling $B = DAD$, where $D = \text{Diag}(r_1, \ldots, r_n)$. Without loss of generality, we assume that $r_1 \leq r_2 \leq \ldots \leq r_n$. Note that since B is doubly stochastic,

$$(6.5) \qquad\qquad r_i = \left(\sum_{j \sim i} r_j \right)^{-1} \,.$$

We will need a few simple lemmas.

Lemma 6.3. *Let $s_1, \ldots, s_n \in [0,1]$ and assume that $\sum_{i=1}^n s_i \geq S$. Then, for any $0 < \gamma < 1$, there exists a subset $I \subseteq [n]$ of cardinality at least $(1 - \gamma) \cdot S$, $s_i \geq \gamma \cdot S/n$ for each $i \in I$.*

Proof. Assume otherwise. Then there are at least $n - (1 - \gamma)S$ elements $s_i < \gamma \cdot (S/n)$. Therefore,

$$\sum_{i=1}^n s_i \leq (1 - \gamma)S + (n - (1 - \gamma)S) \cdot \gamma \frac{S}{n} < S. \quad \square$$

The next lemma quantifies the following intuition: Given a large set $\mathcal{A}$ of indices corresponding to small entries of the scaling matrix, we can find a large set of indices (neighbors of $\mathcal{A}$) corresponding to large entries of the scaling matrix.

Lemma 6.4. *Let $\mathcal{A} \subseteq [n]$ such that $r_i \leq \mu$ for all $i \in \mathcal{A}$. Then, for any $0 < \gamma < 1$, there exists a subset $X \subseteq [n]$, of cardinality at least $(1 - \gamma) \cdot |\mathcal{A}|$, such that for all $j \in X$,*

$$r_j \geq \frac{\gamma}{\mu n}$$

Proof. Denote by $B = (b_{ij})$ the doubly stochastic scaling of A. For $1 \leq i \leq n$, let $s_i = \sum_{j \in \mathcal{A}} b_{ij}$. By the double-stochasticity of B, we have $\sum_{i=1}^n s_i = \sum_{j \in \mathcal{A}} \sum_{i=1}^n b_{ij} = |\mathcal{A}|$.

By Lemma 6.3, there is a set X of indices with $|X| \geq (1 - \gamma) \cdot |\mathcal{A}|$ with $s_i \geq \gamma \cdot \frac{|\mathcal{A}|}{n}$ for each $i \in X$. Let $i \in X$. We have

$$r_i = \frac{s_i}{\sum_{j \in \mathcal{A}, j \sim i} r_j} \geq \gamma \cdot \frac{|\mathcal{A}|/n}{|\mathcal{A}| \cdot \mu} = \frac{\gamma}{\mu n} \,.$$

$\square$

The next lemma is the base of our inductive construction.

Lemma 6.5. *Let* $Q = 1/r_1$. *Then there exists a subset* X *of* $[n]$ *of cardinality at least* $d/2$ *such that for each* $i \in X$,

$$r_i \geq \frac{Q}{2n^2} \,.$$

Proof. By (6.5), $\sum_{i \sim 1} r_i = 1/r_1 = Q$. Therefore, there is at least one index $i_0 \sim 1$ for which $r_{i_0} \geq Q/n$. Let $\mathcal{A}$ be the set of neighbors of i_0. Then, $|\mathcal{A}| \geq d$ and for all $j \in \mathcal{A}$, $r_j \leq n/Q$.

The proof is completed by an application of Lemma 6.4 with $\gamma = 1/2$. $\square$

Lemma 6.6 below will be used for the inductive step.

Lemma 6.6. *Let* $m \geq n$. *Let* X *be a subset of indices such that* $r_i \geq m$ *for each* $i \in X$. *Then*

   (1) $|\partial X| \geq (1 + \kappa) \cdot |X|$;
   (2) *There exists a subset* Z *of indices, disjoint from* X, *of cardinality at least* $(\kappa \cdot |X| - 1)/2$ *such that each* $j \in Z$ *satisfies*

$$r_j \geq \frac{\kappa |X| - 1}{2n^2} \cdot m$$

Proof. Clearly, no two vertices in X are connected in Γ (otherwise, we would have an entry of size at least $m^2 > 1$ after scaling). Therefore X is a set of disconnected vertices, and, since Γ contains a perfect matching, we have $|X| \leq n/2$.

Since X is disconnected, (6.1) and (6.4) imply that

$$|\partial X| \geq (1 + \kappa) \cdot |X|,$$

proving the first claim of the lemma.

Let $Y := \partial X$. We note that $r_i \leq 1/m$ for any $i \in Y$. To show the second claim of the lemma, we will find a subset Z of indices, disjoint with $X \cup Y$, such that $r_i \geq \frac{\kappa |X| - 1}{2n^2} \cdot m$ for all $i \in Z$.

Let $C = (X \cup Y)^{\complement}$. Recall that B is the doubly stochastic scaling of A. Since

$$\sum_{i \in C \cup X \cup Y} \sum_{j \in Y} b_{ij} = \sum_i \sum_{j \in Y} b_{ij} = |Y|,$$

we have

$$\sum_{i\in C, j\in Y} b_{ij} \;=\; |Y| - \sum_{i\in X, j\in Y} b_{ij} - \sum_{i\in Y, j\in Y} b_{ij}$$

$$\geq \;(|Y|-|X|) - \sum_{i\in Y, j\in Y} b_{ij} \;\geq (|Y|-|X|) - \frac{n^2}{m^2}$$

$$\geq \;\kappa \cdot |X| - 1,$$

where in the second inequality we used that $r_i \leq 1/m$ for $i \in Y$.

For $i \in C$, set $s_i = \sum_{j\in Y} b_{ij}$. Then $\sum_{i\in C} s_i \geq \kappa \cdot |X| - 1$. By Lemma 6.3, there is a set of at least $(\kappa \cdot |X| - 1)/2$ indices i, for which

$$s_i \geq \frac{\kappa|X|-1}{2|Y|} \geq \frac{\kappa|X|-1}{2n}.$$

Call this set Z. For each $i \in Z$, we have

$$r_i = \frac{s_i}{\sum_{j\in Y, j\sim i} r_j} \geq \frac{(\kappa|X|-1)/(2n)}{|Y|/m} \geq \frac{\kappa|X|-1}{2n^2} \cdot m,$$

completing the proof of the lemma. $\qquad\square$

We are now ready to perform the inductive procedure proving Proposition 6.1. Let

$$R = \left(\frac{4}{\kappa\alpha} \cdot n\right)^{\frac{c}{\kappa\log(1/\alpha)}},$$

for a sufficiently large c. We will assume that $r_1 < 1/R$, and reach a contradiction.

We use Lemma 6.5 to construct a set X of cardinality at least $d/2$ such that $r_i \geq R/(2n^2)$. for all $i \in X$.

Assuming $m := R/(2n^2) \geq n$, which we may, we can now apply Lemma 6.6 to construct a set Z disjoint from X, of cardinality at least $(\kappa \cdot |X| - 1)/2$ such that

$$r_j \geq \frac{\kappa|X|-1}{2n^2} \cdot m \geq \frac{\kappa\alpha}{4n} \cdot m, \quad \textit{for all } j \in Z.$$

We now define $X_0 := X$, $m_0 := m$, $Z_0 := Z$; and set $X_1 = X_0 \cup Z_0$, $m_1 = \frac{\kappa\alpha}{4n} \cdot m_0$, and apply Lemma 6.6 to X_1 (assuming m_1 is not too small). We continue this process to obtain an increasing sequence of sets $X_0, X_2, \ldots, X_t$. Since

$$n \geq |X_t| \geq (1+\kappa/2)\cdot|X_{t-1}| \geq \ldots \geq (1+\kappa/2)^t \cdot X_0 \geq (1+\kappa/2)^t \cdot \frac{\alpha}{2} \cdot n,$$

the number of steps t is upper bounded by $c_1 \cdot \kappa \log 1/\alpha$, for some absolute constant c_1. On the other hand, if c in the definition of R is

large enough, the number of steps will be larger than that, reaching a contradiction. □

Proof of Proposition 6.2. Assume, for contradiction's sake, that $r_n \geq n^{-\varepsilon/2}$. Since $\sum_{i \sim n} r_i = 1/r_n \leq n^{\varepsilon/2}$, for at least half of neighbors of n holds $r_i \leq 2 \cdot n^{\varepsilon/2-1}$. Let $\mathcal{A}$ be the set of these neighbors. By our assumption on the minimal degree in Γ, we have $|\mathcal{A}| \geq \frac{d}{2}$.

Applying Lemma 6.4 to $\mathcal{A}$ with $\gamma = 1/2$ and $\mu = 2 \cdot n^{\varepsilon/2-1}$ gives a subset X_0 of $[n]$ of cardinality at least $\frac{d}{4}$ such that for all $i \in X_0$ holds $r_i \geq \alpha/8 \cdot n^{-\varepsilon/2}$. This is our induction base.

An inductive step is provided by the following lemma.

Lemma 6.7. *Fix a constant $b \geq 1/\sqrt{\theta}$. Let $X \subseteq [n]$ be such that $|X| \geq \frac{\alpha}{4} \cdot n$ and for any $i \in X$, $r_i \geq b \cdot n^{-\varepsilon/2}$. Then there exists a subset X' of $[n]$ of cardinality at least $\left(1 + \frac{\kappa(1-\kappa)}{2}\right) \cdot |X|$ such that for all $j \in X'$ it holds that $r_j \geq \frac{\alpha^2 \kappa b}{80} \cdot n^{-\varepsilon/2}$.*

Proof. Since $n \geq \sum_{i,j \in X} d_{ij} \geq |E(X,X)| \cdot b^2 n^{-\varepsilon}$, we have
$$|E(X,X)| \leq b^{-2} \cdot n^{1+\varepsilon} \leq \theta \cdot n^{1+\varepsilon}$$
Hence, by our assumptions on the graph Γ, we have $|\partial_s X| \geq (1+\kappa) \cdot |X|$. For each $j \in \partial_s X$ holds
$$r_j \leq \frac{1}{\sum_{i \in X, \ i \sim j} r_i} \leq \frac{1}{(\alpha/10) \cdot |X| \cdot bn^{-\varepsilon/2}} \leq \frac{40}{b\alpha^2} \cdot n^{\varepsilon/2-1}$$

Applying Lemma 6.4 with $\gamma = \kappa/2$ and $\mu = \frac{40}{b\alpha^2} \cdot n^{\varepsilon/2-1}$ to $\mathcal{A} = \partial_s X$, produces a set X' satisfying the requirements of the lemma. □

We now ready to perform the inductive procedure proving Proposition 6.2. Let
$$R = \left(\frac{1}{\kappa} \cdot \log\left(\frac{4}{\alpha}\right)\right)^{\frac{c}{\alpha^2 \kappa}},$$
for a sufficiently large c. We will assume that $\theta > 1/R$, and reach a contradiction.

We start constructing the sequence $\{X_i\}$, starting from the set X_0 constructed above, and applying Lemma 6.7 iteratively. Clearly, we should stop after at most $S = \log_{1+\frac{\kappa(1-\kappa)}{2}}\left(\frac{4}{\alpha}\right)$ steps. However, if c in the definition of R is large enough, we would be able to make more steps than that, reaching a contradiction.

□

We now combine the bound (6.4) on the scaled matrix with Theorems 1.9 and 1.10 to derive Theorem 1.5.

Proof of Theorem 1.5. Recall that $B = DAD$ denotes the doubly stochastic scaling of A and G^{skew} denotes a skew symmetric matrix with independent $N(0, 1)$ entries above the main diagonal. Note that

$$\det(A \odot G^{skew}) = \frac{1}{\det(D)} \det(B_{1/2} \odot G^{skew}),$$

where $B_{1/2}$ denotes the matrix whose entries are the square roots of the entries of B. Therefore, it is enough to consider the concentration for $\det(B_{1/2} \odot G^{skew})$. The proof of Theorem 1.5 now follows by applying Theorems 1.9 and 1.10.   □

7. THE STRONG EXPANSION CONDITION

As noted in the introduction, the strong expansion condition is stronger than the classical vertex expansion condition

$$\forall J \subset [M] \quad |J| \leq M/2 \implies |\partial(J)| \geq \kappa |J|.$$

It might have been desirable to replace the strong expansion property by a weaker and more natural classical vertex expansion condition. Proposition 1.7 from the introduction shows that not only the latter condition is insufficient to guarantee a subexponential error in Barvinok's estimator, but in fact there is an example of a graph G with associated random matrix W that barely misses the strong expansion property, for which Barvinok's estimator yields an exponential error with high probability. We provide here the proof of Proposition 1.7.

Proof of Proposition 1.7. Without loss of generality, assume that $\delta < 1/6$. Let $n \in \mathbb{N}$. Set $m = \lfloor \frac{\delta n}{2} \rfloor$. Define a graph Γ with $M = 2(m + n)$ vertices as follows.

- The vertices in $[n]$ form a clique, which will be called the center.
- Any of the vertices in $[n + 1 : 2(n + m)]$, called peripheral, is connected to all vertices of the center.
- In addition, for $k > n$, the vertices $2k - 1$ and $2k$ are connected to each other.

Vertices $n + 1, \ldots, 2n$

Center: vertices $1, \ldots, n$

Vertices $2n + 1, \ldots, 2(n + m)$
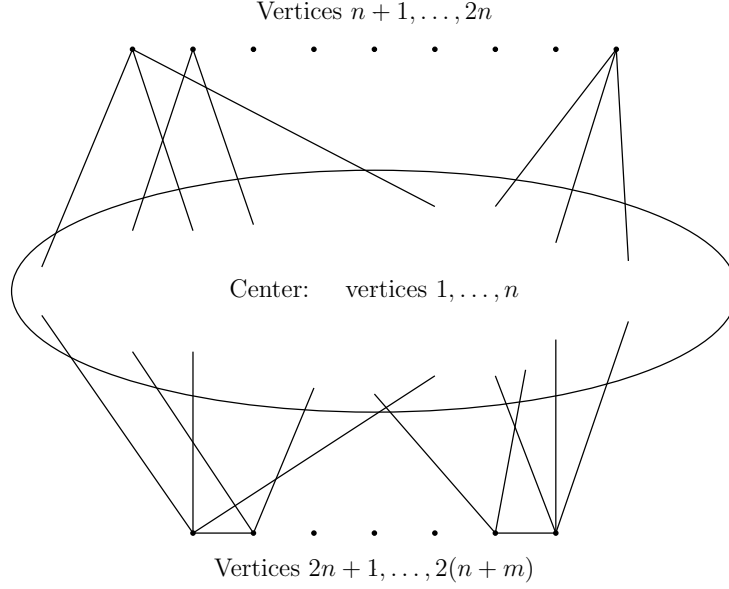
The adjacency matrix of Γ has the block shape

$$
\begin{pmatrix}
\mathbf{Q}_{n \times n} & \mathbf{1}_{n \times n} & \mathbf{1}_{n \times 2m} \\
\mathbf{1}_{n \times n} & \mathbf{0}_{n \times n} & \mathbf{0}_{n \times 2m} \\
& & \boldsymbol{\Delta} \\
\mathbf{1}_{2m \times n} & \mathbf{0}_{2m \times n} & \ddots \\
& & & \boldsymbol{\Delta}
\end{pmatrix} .
$$

Here $\mathbf{Q}_{n \times n}$ is the adjacency matrix of the n-clique, i.e., the matrix with 0 on the main diagonal and 1 everywhere else; $\mathbf{1}_{k \times l}$ is the $k \times l$ matrix whose entries are equal to 1, and Δ is a 2×2 matrix

$$
\Delta = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} .
$$

The right lower block of Γ contains m such matrices Δ on the main diagonal.

The matrix W_Γ has the similar form

$$
W_\Gamma = \begin{pmatrix}
\tilde{\mathbf{Q}}_{n \times n} & \mathbf{G}_{n \times n} & \mathbf{G'}_{n \times 2m} \\
-\mathbf{G}_{n \times n} & \mathbf{0}_{n \times n} & \mathbf{0}_{n \times 2m} \\
& & \tilde{\boldsymbol{\Delta}}_1 \\
-\mathbf{G'}_{2m \times n} & \mathbf{0}_{2m \times n} & \ddots \\
& & & \tilde{\boldsymbol{\Delta}}_m
\end{pmatrix} ,
$$

where $\tilde{\mathbf{Q}}_{n \times n}$ is the $n \times n$ skew-symmetric gaussian matrix; $\mathbf{G}_{n \times n}$, $\mathbf{G}'_{n \times 2m}$ are independent Gaussian matrices, and

$$\tilde{\mathbf{\Delta}}_k = \begin{pmatrix} 0 & g_k \\ -g_k & 0 \end{pmatrix}$$

with independent $N(0, 1)$ random variables $g_1, \dots, g_m$. Recall that

$$\mathbb{E} \det W_\Gamma = \#\text{Matchings}(\Gamma)$$

the number of perfect matchings of the graph Γ. Any vertex from $[n+1 : 2n]$ has to be matched to a vertex from the center, which can be done in $n!$ ways. hence, for $k > n$, any vertex $2k - 1$ has to be matched to its peripheral neighbor $2k$, which can be done in the unique way. Thus,

$$\#\text{Matchings}(\Gamma) = n! > 0.$$

Consider $\det W_\Gamma$. Let $c > 0$ be a constant to be chosen later. A simple pigeonhole argument shows that

$$\det W_\Gamma = F(\mathbf{G}_{n \times n}, \mathbf{G}'_{n \times 2m}) \cdot \prod_{j=1}^{m} g_j^2,$$

where $F(\mathbf{G}_{n \times n}, \mathbf{G}'_{n \times 2m})$ is a homogeneous polynomial of degree $2n$ of entries of $\mathbf{G}_{n \times n}$ and $\mathbf{G}'_{n \times 2m}$). Hence, for $\alpha = \frac{4}{\delta} + 1$, we have

$$\mathbb{P}\left(\frac{\det(W_\Gamma)}{\mathbb{E}\det(W_\Gamma)} \geq e^{-cM} \right) = \mathbb{P}\left(\frac{\det(W_\Gamma)}{\mathbb{E}\det(W_\Gamma)} \geq \exp(-c\alpha m) \right)$$

$$\leq \mathbb{P}\left(\frac{F(\mathbf{G}_{n \times n}, \mathbf{G}'_{n \times 2m})}{\mathbb{E}F(\mathbf{G}_{n \times n}, \mathbf{G}'_{n \times 2m})} \geq \exp(c\alpha m) \right)$$

$$+ \mathbb{P}\left(\frac{\prod_{j=1}^{m} g_j^2}{\mathbb{E}\prod_{j=1}^{m} g_j^2} \geq \exp(-2c\alpha m) \right)$$

The first term above is smaller than $\exp(-c\alpha m)$ by the Chebyshev inequality. The second term also does not exceed $\exp(-c'm)$ if the constant c is chosen small enough.

This proves the part of the proposition related to the error of the Barvinok estimator.

It remains to check that the condition (1.11) is satisfied. Let $J \subset [M]$ be a set of cardinality $|J| \leq M/2$. If J contains a vertex from the center, then $|\text{Con}(J)| = 1$ and $\partial(J) = [M] \setminus J$, so condition (1.11) holds.

Assume that $J \cap [n] = \emptyset$. Then $|\text{Con}(J)| \leq m + n = M/2$. Also, $\partial(J) \supset [n]$, so

$$|\partial(J)| \geq n \geq \frac{1}{1 + \delta/2} \cdot \frac{M}{2}.$$

Therefore, since $\delta < 1/6$,

$$|\partial(J)| - (1 - \delta)|\text{Con}(J)| \geq \left(\frac{1}{1 + \delta/2} - (1 - \delta) \right) \cdot \frac{M}{2} \geq \frac{\delta}{8} \cdot \frac{M}{2} \geq \kappa \cdot |J|$$

if we choose $\kappa = \delta/8$. This completes the proof of the proposition. $\quad\square$

References

[1] O. Ajanki, L. Erdős and T. Krüger, *Local semicircle law with imprimitive variance matrix*, Arxiv:1311.2016 (2013).

[2] A. Barvinok, *Polynomial time algorithms to approximate permanents and mixed discriminants within a simply exponential factor*, Random Structures Algorithms **14** (1999), pp. 29–61.

[3] A. Barvinok and A. Samorodnitsky, *Computing the partition function for perfect matchings in a hypergraph*, Combin. Probab. Comput. **20** (2011), pp. 815–835.

[4] A. Barvinok and A. Samorodnitsky, *Random weighting, asymptotic counting, and inverse isoperimetry*, Israel J. Math. **158** (2007), pp. 159–191.

[5] M. Bayati, D. Gamarnik, D. Katz, C. Nair and P. Tetali, *Simple deterministic approximation algorithms for counting matchings*, STOC'07 - Proceedings of the 39th Annual ACM Symposium on Theory of Computing, pp. 122–127, ACM, New York, 2007.

[6] L. M. Bregman, *Some properties of nonnegative matrices and their permanents*, Soviet Math. Dokl. **211** (1973), pp. 945–949.

[7] L. Erdős, A. Knowles, H.-T. Yau and J. Yin, *The local semicircle law for a general class of random matrices*, Elec. J. Probab. **18** (2013), paper 59.

[8] S. Friedland, B. Rider and O. Zeitouni, *concentration of permanent estimators for certain large matrices*, Annals Appl. Prob **14** (2004), pp. 1359–1576.

[9] C. D. Godsil and I. Gutman, *on the matching polynomial of a graph*, in *Algebraic methods in graph theory I-II* (L. Lóvasz and V. T. Sós, eds.), North-Holland, Amsterdam (1981), pp. 67–83.

[10] A. Guionnet and O. Zeitouni, *Concentration of the spectral measure for large matrices*, Elec. Comm. Probab. **5** (2000), pp. 119–136.

[11] M. Jerrum and A. Sinclair, *Approximating the permanent*, SIAM J. Comput. **18** (1989), pp. 1149–1178.

[12] M. Jerrum, A. Sinclair and E. Vigoda, *A polynomial-time approximation algorithm for the permanent of a matrix with non-negative entries*, J. ACM **51** (2004), pp. 671–697.

[13] M. Ledoux, *The concentration of measure phenomenon*, American Math. Soc. (2001).

[14] N. Linial, A. Samorodnitsky and A. Wigderson, *A deterministic strongly polynomial algorithm for matrix scaling and approximate permanents*, Combinatorica **20** (2000), pp. 545–568.

[15] H. Minc, *Permanents*, Addison-Wesley, Reading, MA

[16] M. Rudelson, R. Vershynin, *The Littlewood-Offord Problem and invertibility of random matrices*. Adv. Math **218** (2008), pp. 600–633.

[17] M. Rudelson, R. Vershynin, *The smallest singular value of a random rectangular matrix*. Comm. Pure Appl. Math. **62** (2009), pp. 1707–1739.

[18] M. Rudelson, O. Zeitouni, *Singular values of Gaussian matrices and permanent estimators*, to appear, RSA (2014). arXiv:1301.6268.

[19] S. Szarek, *Spaces with large distance to ℓ_∞^n and random matrices.* Amer. J. Math. **112** (1990), pp. 899–942.

[20] L. Valiant, *The complexity of evaluating the permanent*, Theoret. Comput. Sci. **8** (1979), pp. 189–201.