# Smoothed Analysis of the Komlós Conjecture: Rademacher Noise

Elad Aigner-Horev[a] Dan Hefetz[b] Michael Trushkin[c]

Submitted: Jul 20, 2024; Accepted: Feb 3, 2025; Published: Mar 28, 2025


**Abstract**

The *discrepancy* of a matrix $M \in \mathbb{R}^{d\times n}$ is given by $\mathrm{DISC}(M) := \min_{\mathbf{x}\in\{-1,1\}^n} \|M\mathbf{x}\|_\infty$. An outstanding conjecture, attributed to Komlós, stipulates that $\mathrm{DISC}(M) = O(1)$, whenever $M$ is a Komlós matrix, that is, whenever every column of $M$ lies within the unit sphere. Our main result asserts that $\mathrm{DISC}(M + R/\sqrt{d}) = O(d^{-1/2})$ holds asymptotically almost surely, whenever $M \in \mathbb{R}^{d\times n}$ is Komlós, $R \in \mathbb{R}^{d\times n}$ is a Rademacher random matrix, $d = \omega(1)$, and $n = \omega(d\log d)$. The factor $d^{-1/2}$ normalising $R$ is essentially best possible and the dependency between $n$ and $d$ is asymptotically best possible. Our main source of inspiration is a result by Bansal, Jiang, Meka, Singla, and Sinha (ICALP 2022). They obtained an assertion similar to the one above in the case that the smoothing matrix is Gaussian. They asked whether their result can be attained with the optimal dependency $n = \omega(d\log d)$ in the case of Bernoulli random noise or any other types of discretely distributed noise; the latter types being more conducive for Smoothed Analysis in other discrepancy settings such as the Beck-Fiala problem. For Bernoulli noise, their method works if $n = \omega(d^2)$. In the case of Rademacher noise, we answer the question posed by Bansal, Jiang, Meka, Singla, and Sinha. Our proof builds upon their approach in a strong way and provides a discrete version of the latter.

**Mathematics Subject Classifications:** 05C88, 05C89

## 1 Introduction

The *discrepancy* of a matrix $M \in \mathbb{R}^{d\times n}$ is given by $\mathrm{DISC}(M) := \min_{\mathbf{x}\in\{-1,1\}^n} \|M\mathbf{x}\|_\infty$. A celebrated result in this venue is the so-called "six standard deviations" result, put forth by Spencer [45], asserting that if $\|M\|_\infty \leqslant 1$ and $d = n$, then $\mathrm{DISC}(M) \leqslant 6\sqrt{n}$. More generally, if $d \geqslant n$, then $\mathrm{DISC}(M) = O\left(\sqrt{n\log(2d/n)}\right)$ is known to hold [15, 33, 39, 44]. Spencer's

[a] School of Computer Science, Ariel University, Israel (horev@ariel.ac.il).
[b] School of Computer Science, Ariel University, Israel (danhe@ariel.ac.il).
[c] School of Computer Science, Ariel University, Israel (michaelt@ariel.ac.il).

result is essentially tight as $n \times n$ matrices $M$ satisfying $\mathrm{DISC}(M) = \Omega(\sqrt{n})$ are known to exist [26].

An outstanding conjecture in Discrepancy Theory, attributed to Komlós, stipulates that $\mathrm{DISC}(M) = O(1)$ holds, whenever $M \in \mathbb{R}^{d\times n}$ has each of its columns $\mathbf{v}$ satisfying $\|\mathbf{v}\|_2 \leqslant 1$; we refer to the latter as a *Komlós matrix*[1]. *Dimension-free* (i.e., constant) bounds on the discrepancy of Komós matrices are of special interest, in particular, because it is $NP$-hard to distinguish between Komlós matrices having zero discrepancy and those having discrepancy one [26].

Given a hypergraph $H$, taking $M = M_H$ to be its $e(H) \times v(H)$ incidence matrix retrieves the well-known (see, e.g., [18, 27]) notion of *combinatorial discrepancy*, given by

$$\mathrm{DISC}(H) := \min_{\chi} \max_{e\in E(H)} \left|\sum_{v\in e} \chi(v)\right|,$$

where the minimisation ranges over all mappings $\chi : V(H) \to \{-1, 1\}$. Beck and Fiala [17] proved that if $H$ has the property that each of its vertices lies in at most $t$ edges, i.e., each column $\mathbf{v}$ of $M_H$ satisfies $\|\mathbf{v}\|_2 \leqslant \sqrt{t}$, then $\mathrm{DISC}(H) \leqslant 2t-1$, and conjectured that $\mathrm{DISC}(H) = O(\sqrt{t})$, in fact, holds in this case. The Beck-Fiala conjecture follows as a special case of the Komlós conjecture upon rescaling by $\sqrt{t}$. The best known upper bounds for the conjectures put forth by Komlós and by Beck-Fiala are $O\left(\sqrt{\log n}\right)$ and $O\left(\sqrt{t \log n}\right)$, respectively, both obtained by Banaszczyk [14] in 1998[2]. Despite this partial progress, it seems that these two conjectures are out of reach of current techniques; consequently, the investigation of these conjectures in more hospitable settings, so to speak, is well-justified.

One line of research that has attracted much attention of late calls for the determination of $\mathrm{DISC}(M)$ whenever $M$ is a random matrix; in this line of research one is interested in the so-called *average-case* discrepancy or the discrepancy of *typical* matrices, where 'typical' depends on the specific distribution chosen for $M$. In this realm, we further distinguish between two strands of study; the first pertains to Gaussian matrices[3] and the second deals with discrete random matrices.

For standard Gaussian matrices $M \in \mathbb{R}^{d\times n}$, the estimate $\mathrm{DISC}(M) = \Theta\left(2^{-n/d}\sqrt{n}\right)$ holds asymptotically almost surely (a.a.s. hereafter) for a wide range of values of $d$ and $n$; in particular $\mathrm{DISC}(M) = O(1)$ holds as soon as $n \geqslant Cd \log d$, where $C > 0$ is an appropriate constant. This without $M$ necessarily being Komlós[4]. The case $d = O(1)$ of the above equality was settled by Costello [28]. Meka, Rigollet, and Turner [41] extended the result of Costello by

[1]Komlós' restriction on the matrix is more stringent than that of Spencer.

[2]For the Beck-Fiala conjecture, see also [**?**] for the currently best bound which is independent of $n$.

[3]Matrices with each entry an i.i.d. copy of $\mathcal{N}(\mu, \sigma^2)$; if $\mu = 0$ and $\sigma = 1$, then the matrix is called a *standard* Gaussian matrix.

[4]Producing Komlós matrices from standard Gaussian matrices $M \in \mathbb{R}^{d\times n}$ is straightforward. To see this, recall that a column $\mathbf{v}$ of such a matrix satisfies $\|\mathbf{v}\|_2 \approx \sqrt{d}$ a.a.s. (see, e.g., [51, Theorem 3.1.1]); this property extends to all columns of $M$ following an appropriate union-bound calculation. Hence, normalising all columns by (essentially) $\sqrt{d}$ produces a Komlós matrix asymptotically almost surely.

allowing $\omega(1) = d = o(n)$. In fact, their result accommodates any (matrix entry) distribution whose density function $f$ is symmetric, has a fourth moment, and is square-integrable. The regime $d = \Theta(n)$ was studied in [1, 12, 25, 42].

Proceeding to discrete random matrices, given $d \geqslant n \geqslant t$, Ezra and Lovett [31] proved that $\mathrm{DISC}(M) = O\left(\sqrt{t \log t}\right)$ holds with probability at least $1 - \exp(-\Omega(t))$, whenever each column of $M \in \{0,1\}^{d \times n}$ is sampled independently and uniformly at random from all 0/1-vectors containing precisely $t$ non-zero entries. Normalising each column of $M$ by its 2-norm (i.e. $\sqrt{t}$) produces a Komlós matrix, thus implying that Komós matrices $N$ produced in this fashion satisfy $\mathrm{DISC}(N) = O\left(\sqrt{\log t}\right)$ with probability at least $1 - \exp(-\Omega(t))$. For sparser such matrices satisfying $d \geqslant t$ and $n \gg d^t$, Ezra and Lovett proved that $\mathrm{DISC}(M) = O(1)$ holds asymptotically almost surely. For Bernoulli matrices[5] $M \in \mathbb{R}^{d \times n}$, Altschuler and Niles-Weed [11] proved that $\mathrm{DISC}(M) \leqslant 1$ holds a.a.s. for any $p := p(n)$, whenever $n \geqslant Cd \log d$, where $C > 0$ is an absolute constant[6]; their result is tight in terms of the lower bound on $n$. Moreover, their bound on the discrepancy is also best possible as any binary matrix that has a row with an odd number of 1's has discrepancy at least one.

Given a *seed* matrix $M \in \mathbb{R}^{d \times n}$ as well as a distribution $\mathcal{R}_{d \times n}$, set over $\mathbb{R}^{d \times n}$, we refer to the (random) matrix $M + R$ with $R \sim \mathcal{R}_{d \times n}$ as a *random perturbation* of $M$. Following the aforementioned results pertaining to the discrepancy of fully random matrices, the study of the discrepancy of randomly perturbed ones is the next natural step. The study of the effect of random *noise* is widespread in Mathematics and Computer Science. Spielman and Teng [47] coined the term *smoothed analysis* to indicate the analysis of algorithms executed on randomly perturbed inputs. In high dimensional probability (see, e.g., [51]), the study of randomly perturbed matrices dates back to the works of Tao and Vu [50, 48, 49]. In combinatorics, the study of randomly perturbed (hyper)graphs has witnessed a burst of activity in recent years; see, e.g., [2, 3, 4, 6, 5, 7, 8, 9, 10, 13, 19, 20, 21, 24, 23, 29, 32, 37, 36, 38, 40].

The main source of inspiration for our work is a result by Bansal, Jiang, Meka, Singla, and Sinha [16] who established the first ever perturbed/smoothed version of the Komlós conjecture. They proved that $\mathrm{DISC}(M+R) \leqslant \frac{1}{\mathrm{poly}(d)}$ holds a.a.s. whenever $M \in \mathbb{R}^{d \times n}$ is a Komlós matrix, $R \in \mathbb{R}^{d \times n}$ is a matrix whose entries are i.i.d. copies of $\frac{\sigma}{\sqrt{d}}\mathcal{N}(0,1)$ and $n = \omega(d \log d) \cdot \sigma^{-4/3}$. In [16, Section 3], Bansal, Jiang, Meka, Singla, and Sinha ask whether analogous results can be proved if instead of Gaussian noise one uses discrete noise such as the Bernoulli distribution or some other natural discrete distribution. The interest in discrete noise models is reasoned in [16] as being more conducive for Smoothed Analysis in other discrepancy settings such as the Beck-Fiala problem. Bansal, Jiang, Meka, Singla, and Sinha [16, Section 3] note that they are able to prove the required results if the smoothing noise has the Bernoulli distribution and $n = \omega(d^2)$. Attaining the optimal dependency $n = \omega(d \log d)$ for discrete noise models is then of interest and seems to require additional tools.

[5]Each entry is an independent copy of $\mathrm{Ber}(p)$ for $p := p(n,d)$.

[6]Discrepancy of Poisson matrices is also studied in [11]; Bernoulli matrices are also studied in [35, 43].

### 1.1 Our contribution

A random variable $X$ is said to be *Rademacher* if $X$ assumes the values $-1$ and 1, each with probability 1/2. A matrix $R \in \mathbb{R}^{d\times n}$ is said to form a *Rademacher matrix* if its entries are independent Rademacher random variables. Our main result reads as follows.

**Theorem 1.** *Let $d = \omega(1)$ and $n = \omega(d \log d)$ be integers. Then,* $\mathrm{DISC}(M + R/\sqrt{d}) \leqslant 8d^{-1/2}$ *holds a.a.s. whenever $M \in \mathbb{R}^{d\times n}$ is a Komlós matrix and $R \in \mathbb{R}^{d\times n}$ is a Rademacher matrix.*

This resolves the aforementioned question of Bansal, Jiang, Meka, Singla, and Sinha [16, Section 3] in the case that the smoothing noise has the Rademacher distribution. In the case that the noise has the Bernoulli distribution, the aforementioned dependency $n = \omega(d^2)$, stated in [16, Section 3], is the state of the art.

*Remark* 2. **Normalisation factor - lower bound.** In Theorem 1, the Rademacher matrix $R$ is normalised by a $\sqrt{d}$ factor. This normalisation factor is warranted. Indeed, requiring that $\|\mathbf{v}\|_2 \leqslant 1$ holds for every column $\mathbf{v}$ of the random perturbation is a natural constraint to impose, for such a restriction guarantees that the columns of the perturbation do not dominate the columns of $M$. Writing $k := k(d)$ to denote the normalisation factor and letting $\mathbf{v}$ be any column vector of $R/k$, we see that $1 \geqslant \|\mathbf{v}\|_2^2 = \sum_{i=1}^{d} \frac{1}{k^2} = \frac{d}{k^2}$ implies $k \geqslant \sqrt{d}$.

*Remark* 3. **Normalisation factor - upper bound.** Let $k$ be as defined in Remark 2. Enlarging $k$ is of interest as this reduces the dominance of the random perturbation further, allowing one to come ever closer to Komlós' conjecture. Alas, in the setting of Theorem 1, there is an upper bound on the normalisation factor $k$. To see this, note that given $k$ and a discrepancy bound $\Delta$, the stipulation that $\mathrm{DISC}(M + R/k) \leqslant \Delta$ is equivalent to requiring the existence of a vector $\mathbf{x} \in \{-1, 1\}^n$ for which

$$(R\mathbf{x})_i \in [-k(M\mathbf{x})_i - k\Delta, -k(M\mathbf{x})_i + k\Delta] \tag{1}$$

holds for every $i \in [d]$. Given $\mathbf{x} \in \{-1, 1\}^n$ and $i \in [d]$, the term $(R\mathbf{x})_i$ has the same distribution as the sum $\sum_{i=1}^{n} r_i$, whose summands are independent Rademacher random variables. As such, $(R\mathbf{x})_i \in [-\omega(\sqrt{n}), \omega(\sqrt{n})]$ asymptotically almost surely. Consequently, a prerequisite for (1) holding a.a.s. is that

$$[-k(M\mathbf{x})_i - k\Delta, -k(M\mathbf{x})_i + k\Delta] \cap [-\omega(\sqrt{n}), \omega(\sqrt{n})] \neq \emptyset$$

holds for every $i \in [d]$. Assuming that $\Delta$ is relatively small (as one naturally aims to have), the latter amounts to essentially requiring that $k \leqslant \sqrt{n}\|M\mathbf{x}\|_\infty^{-1}$. The smaller the value of $\|M\mathbf{x}\|_\infty$ we obtain, the less restrictive on $k$ this inequality becomes. In our current state of knowledge, the best we can ensure are vectors $\mathbf{x} \in \{-1, 1\}^n$ for which $\|M\mathbf{x}\|_\infty = O(\sqrt{\log d})$ (see Lemma 7(ii)). Such a vector then yields the upper bound $k = O\left(\sqrt{n/\log d}\right)$. It follows that for $n = \omega(d \log d)$ (as in the premise of Theorem 1), taking $k$ to be roughly $\sqrt{d}$ is essentially best possible, subject to the aforementioned state of the art worst-case bounds in this venue.

*Remark* 4. **Dependence between $n$ and $d$.** The requirement $n = \omega(d \log d)$ appearing in Theorem 1 is asymptotically best possible. To see this, take $M$ to be the zero matrix (which is Komlós) and note that it suffices to prove that $\mathrm{DISC}(R) = O(1)$ mandates $n = \omega(d \log d)$. To this end, fix an arbitrary vector $\mathbf{x} \in \{-1, 1\}^n$. Given any constant $C > 0$ and a row $\mathbf{r}$ of the Rademacher matrix $R \in \{-1, 1\}^{d \times n}$, we may write

$$\mathbb{P}\left[|\langle \mathbf{r}, \mathbf{x} \rangle| \leqslant C\right] = \sum_{k=-C/2}^{C/2} \frac{\binom{n}{n/2+k}}{2^n} = O\left(n^{-1/2}\right),$$

where the first equality holds since $\langle \mathbf{r}, \mathbf{x} \rangle$ has the same distribution as a sum of i.i.d. Rademacher random variables, and the second equality is supported by Proposition 11 below. It follows by the independence of the entries of $R$ that $\mathbb{P}[\|R\mathbf{x}\|_\infty \leqslant C] = O\left(n^{-d/2}\right)$. Call a vector $\mathbf{x} \in \{-1, 1\}^n$ *good* if $\|R\mathbf{x}\|_\infty \leqslant C$. Then

$$\mathbb{E}[\text{number of good vectors}] = O\left(2^n n^{-d/2}\right) = O\left(\exp(n - d \log n/2)\right).$$

If $n = o(d \log d)$, then the above expectation vanishes and it thus follows by Markov's inequality that a.a.s. no good vectors exist.

### 1.2 Our approach

A natural approach towards proving Theorem 1 is to use the CLT in conjunction with the work of Bansal, Jiang, Meka, Singla, and Sinha [16] for the case of Gaussian noise. Our attempts to employ this approach stalled at the sub-optimal dependency $n = \omega(d^2)$; the same bound attained in [16] for Bernoulli noise. To break the $n = \omega(d^2)$ barrier and reach the optimal dependency $n = \omega(d \log d)$, we utilise and build upon the framework of Bansal, Jiang, Meka, Singla, and Sinha [16] in a strong way. In that, we provide a discretisation of their argument without appealing to the CLT. This calls for a meticulous counting argument whose execution requires new ideas and various adaptations of the arguments of [16].

In the Gaussian case, Bansal, Jiang, Meka, Singla, and Sinha [16] employ the so-called *weighted second moment method* through which they equip $\{-1, 1\}^n$ with an appropriate distribution (see [16, Lemma 2.1] and also Lemma 7) from which a vector $\mathbf{x} \in \{-1, 1\}^n$ satisfying $\|(M + R)\mathbf{x}\|_\infty \leqslant 1/\mathrm{poly}(d)$ is identified, where here $R$ is a (scaled) conformal[7] Gaussian matrix. Writing $S$ for the number of vectors $\mathbf{x}$ satisfying the above bound, they employ the Paley-Zygmund inequality (see (10)) in order to prove that $\mathbb{P}[S > 0] \geqslant 1 - o(1)$. The bulk of the argument consists of establishing that $\mathbb{E}_R[S^2] \leqslant (1 + o(1))\, \mathbb{E}_R[S]^2$. This is the core of our proof of Theorem 1 as well; this also provides a very crude high level description of the approach of [16] which we follow.

Executing the above approach with Rademacher noise summons various challenges. Roughly put, in [16] one, for instance, encounters the need to estimate the probability that a Gaussian

[7]Throughout, we use the term *conformal* to mean that an associated algebraic structure has the adequate dimensions set for it; the latter, however, are seen as distracting and are thus omitted.

vector lies, say, within some rectangular region. In the Gaussian case, this can be handled through a nontrivial estimation (as seen in [16]) of a certain integral involving the probability density function of an adequate multi-dimensional Gaussian distribution. A similar situation in the Rademacher case, becomes rather involved, especially in view of the need to not lose track of the optimal dependency $n = \omega(d \log d)$; these types of problems require a careful counting argument, to which we gain initial access through Proposition 11, providing us with a rather tight estimation for certain binomial coefficients encountered throughout. The need to take on such discretised estimations abound throughout our proof and these collectively entail that various additional subtle discretised adjustments and adaptations be made to several ingredients seen in the work of [16]; effect of which is not clear from the offset. For example, we require an adaption of the original aforementioned distribution defined in [16, Lemma 2.1] (see Lemma 9).

The rather meticulous nature of our core counting argument makes it so that supplying additional more concrete examples of the work done in our proof entails the reconstruction of somewhat vast supporting sceneries required for the context. We thus refrain from providing further details at this stage and refer the reader to Section 3 for an accurate account.

## 2 Preliminaries

This section is divided into two subsections, both containing auxiliary results facilitating our proof of Theorem 1.

*Remark* 5. Throughout this section we encounter binomial coefficients of the form $\binom{n}{n/2+t}$, where $n \in \mathbb{N}$ is even and $t \in \mathbb{Z}$. Owing to the symmetry $\binom{n}{n/2+t} = \binom{n}{n/2-t}$, whenever it is convenient, we assume that $t \geqslant 0$.

### 2.1 Key tools

Our overall goal in Theorem 1 is to prove the existence of a vector $\mathbf{x} \in \{-1,1\}^n$ for which a.a.s. $\|(M + R/\sqrt{d})\mathbf{x}\|_\infty \leqslant 8d^{-1/2}$ holds. In order to do so we follow the core innovative technique put forth by Bansal, Jiang, Meka, Singla, and Sinha [16] and sample the vectors of $\{-1,1\}^n$ according to an adaptation of a distribution $\mathcal{D} := \mathcal{D}_n$, called the *truncated Gram-Schmidt distribution*, defined below in Lemma 7.

A real random variable $X$ is said to be *$\alpha$-subgaussian*[8] if it satisfies $\mathbb{P}[|X| \geqslant t] \leqslant 2\exp(-(t/\alpha)^2)$ for every $t > 0$. A random vector $\mathbf{x} \in \mathbb{R}^n$ is said to be *$\alpha$-subgaussian* if $\langle \mathbf{x}, \mathbf{y} \rangle$ is $\alpha$-subgaussian for every $\mathbf{y} \in \mathbb{S}^{n-1}$, see, e.g., [51, Definition 3.4.1]. The following is one of the main results of [34].

**Theorem 6.** [34] *Let $\mathbf{v}^{(1)}, \dots, \mathbf{v}^{(n)} \in \mathbb{R}^m$ satisfy $\|\mathbf{v}^{(i)}\|_2 \leqslant 1$ for every $i \in [n]$. Applying*

[8]Subgaussian random variables admit several equivalent characterisations; see, e.g., [51, Proposition 2.5.2] for details.

*the Gram-Schmidt walk sampling algorithm*[9] *over the given vectors outputs a random vector* $\mathbf{x} \in \{-1,1\}^n$ *such that the vector* $\sum_{i=1}^n \mathbf{x}_i \mathbf{v}^{(i)}$ *is* 1*-subgaussian.*

The distribution (implicitly) defined in Theorem 6 is *truncated* in [16] so as to produce the following distribution over the vectors in $\{-1,1\}^n$.

**Lemma 7.** [16, Lemma 2.1] *Let* $M \in \mathbb{R}^{d\times n}$ *be a Komlós matrix. Then, there exists a constant* $C_7 > 0$ *as well as a distribution* $\mathcal{D} := \mathcal{D}_n$*, set over the vectors in* $\{-1,1\}^n$*, such that the following three properties hold simultaneously.*

- *(i)* $\|M\mathbf{x}\|_2 \in [r-\mathcal{T}, r+\mathcal{T}]$ *holds for every* $\mathbf{x} \in \operatorname{Supp} \mathcal{D}$*, where* $r = O(\sqrt{d})$ *and* $\mathcal{T} = d^{-C'}$ *for some constant* $C' > 1$*.*
- *(ii)* $\|M\mathbf{x}\|_\infty = O\left(\sqrt{\log d}\right)$ *holds for every* $\mathbf{x} \in \operatorname{Supp} \mathcal{D}$*.*
- *(iii)* $\mathbb{P}_{\mathbf{x}\sim\mathcal{D}}\left[|\langle \mathbf{x}, \mathbf{u}\rangle| \geqslant t\right] \leqslant d^{C_7}\exp(-t^2/8)$ *and* $\mathbb{P}_{\mathbf{x}\sim\mathcal{D}}\left[|\langle M\mathbf{x}, \mathbf{v}\rangle| \geqslant t\right] \leqslant d^{C_7}\exp(-t^2/8)$ *both hold whenever* $\mathbf{u} \in \mathbb{S}^{n-1}$*,* $\mathbf{v} \in \mathbb{S}^{d-1}$*, and* $t > 0$*.*

*Remark* 8. Part (ii) of Lemma 7 is not stated in [16]. Its proof being essentially the same as the proof of Lemma 7(i) in [16] is thus omitted.

We require certain extensions of the distribution $\mathcal{D}$. These extensions form a key difference between our discrete setting and the Gaussian case seen in the work of [16]. The contribution of these extensions is not restricted to a single specific point in our arguments and they are thus not so easy to motivate at this stage. Roughly put, these extensions allow us to obtain more control over the structure of the members found in the support of (a deterministic alteration of) the distribution $\mathcal{D}$ used in the Gaussian case. For instance, in Claim 17, the use of such extensions allows us to generate a distribution which is akin to $\mathcal{D}$ and has the added feature that all members in its support have an even number of entries equal to one.

Given a non-negative integer $k$ and an injective mapping $\varphi : \{-1,1\}^n \to \{-1,1\}^{n+k}$, a distribution $\mathcal{S}$ over $\{-1,1\}^{n+k}$ is said to be a *deterministic* $\varphi$*-extension* of $\mathcal{D}_n$ if the latter can be obtained by first sampling a vector $\mathbf{x} \sim \mathcal{D}_n$ and then applying $\varphi$ to $\mathbf{x}$. If the specific nature of the mapping $\varphi$ is inconsequential, then we simply say that $\mathcal{S}$ forms a *deterministic extension* of $\mathcal{D}_n$. The following is an adaptation of Lemma 7, applicable to $\mathcal{S}$.

**Lemma 9.** *Let* $i \in \{1,2\}$ *and let* $M \in \mathbb{R}^{d\times n}$ *be a Komlós matrix whose last* $i$ *columns form the zero vector* $\mathbf{0}$*. Let* $\mathcal{S}$ *be a distribution over* $\{-1,1\}^n$ *which forms a deterministic extension of* $\mathcal{D}_{n-i}$*. Then, there exists a constant* $C_9 > 0$ *such that the following three properties hold simultaneously.*

- *(i)* $\|M\mathbf{x}\|_2 \in [r-\mathcal{T}, r+\mathcal{T}]$ *holds for every* $\mathbf{x} \in \operatorname{Supp} \mathcal{S}$*, where* $r = O(\sqrt{d})$ *and* $\mathcal{T} = d^{-C'}$ *for some constant* $C' > 1$*.*
- *(ii)* $\|M\mathbf{x}\|_\infty = O\left(\sqrt{\log d}\right)$ *holds for every* $\mathbf{x} \in \operatorname{Supp} \mathcal{S}$*.*

[9] See [34] for details.

(*iii*) $\mathbb{P}_{\mathbf{x}\sim\mathcal{S}}\left[|\langle \mathbf{x},\mathbf{u}\rangle| \geqslant t\right] \leqslant d^{C_9}\exp(-t^2/9)$ *and* $\mathbb{P}_{\mathbf{x}\sim\mathcal{S}}\left[|\langle M\mathbf{x},\mathbf{v}\rangle| \geqslant t\right] \leqslant d^{C_9}\exp(-t^2/8)$ *both hold whenever* $\mathbf{u}\in\mathbb{S}^{n-1}$, $\mathbf{v}\in\mathbb{S}^{d-1}$, *and* $t>0$.

*Proof.* Since the last $i$ columns of $M$ are $\mathbf{0}$, parts (i) and (ii) and the fact that $\mathbb{P}_{\mathbf{x}\sim\mathcal{S}}\left[|\langle M\mathbf{x},\mathbf{v}\rangle| \geqslant t\right] \leqslant d^{C_9}\exp(-t^2/8)$ holds for every $\mathbf{v}\in\mathbb{S}^{d-1}$ and every $t>0$, are immediate corollaries of their counterparts in Lemma 7.

Fix $t>0$ and an arbitrary vector $\mathbf{u}\in\mathbb{S}^{n-1}$. For every vector $\mathbf{v}\in\mathbb{R}^n$, let $\mathbf{v}^{(i)}\in\mathbb{R}^{n-i}$ be the vector consisting of the first $n-i$ coordinates of $\mathbf{v}$. Then

$$\begin{aligned}\mathbb{P}_{\mathbf{x}\sim\mathcal{S}}\left[|\langle \mathbf{x},\mathbf{u}\rangle| \geqslant t\right] &\leqslant \mathbb{P}_{\mathbf{x}^{(i)}\sim\mathcal{D}_{n-i}}\left[\left|\left\langle \mathbf{x}^{(i)},\mathbf{u}^{(i)}\right\rangle\right| \geqslant t-2\right]\\ &= \mathbb{P}_{\mathbf{x}^{(i)}\sim\mathcal{D}_{n-i}}\left[\left|\left\langle \mathbf{x}^{(i)},\|\mathbf{u}^{(i)}\|_2^{-1}\mathbf{u}^{(i)}\right\rangle\right| \geqslant (t-2)\|\mathbf{u}^{(i)}\|_2^{-1}\right]\\ &\leqslant d^{C_7+1}\exp\left(-(t-2)^2/8\right)\\ &\leqslant d^{C_9}\exp\left(-t^2/9\right),\end{aligned}$$

where the first inequality holds since $\mathbf{x}\in\{-1,1\}^n$ and $\mathbf{u}\in\mathbb{S}^{n-1}$, and the second inequality holds by Lemma 7(iii) and since $\|\mathbf{u}^{(i)}\|_2\leqslant\|\mathbf{u}\|_2=1$ (multiplying by $d$ is used to circumvent the case $t\leqslant 2$; additionally, if $\|\mathbf{u}^{(i)}\|_2=0$, then the claim is trivial). □

Given two vectors $\mathbf{x},\mathbf{y}\in\{-1,1\}^n$, let $\varepsilon:=\varepsilon(\mathbf{x},\mathbf{y})=\frac{\langle\mathbf{x},\mathbf{y}\rangle}{n}$, let $\mathrm{Diff}(\mathbf{x},\mathbf{y})=\{i\in[n]:\mathbf{x}_i\neq\mathbf{y}_i\}$, and let $\alpha:=\alpha(\mathbf{x},\mathbf{y})=1-\frac{|\mathrm{Diff}(\mathbf{x},\mathbf{y})|}{n}$. Note that $|\mathrm{Diff}(\mathbf{x},\mathbf{y})|$ is the Hamming distance between $\mathbf{x}$ and $\mathbf{y}$, and $\alpha(\mathbf{x},\mathbf{y})n$ is the number of indices over which these two vectors coincide. The following result presents simple but useful relations between these parameters.

**Claim 10.** *Let* $\mathbf{x},\mathbf{y}\in\{-1,1\}^n$ *and let* $\varepsilon:=\varepsilon(\mathbf{x},\mathbf{y})$ *and* $\alpha:=\alpha(\mathbf{x},\mathbf{y})$ *be as above. Then*

(*a*) $\alpha=\frac{1+\varepsilon}{2}$;

(*b*) $\frac{1}{\alpha(1-\alpha)}=\frac{4}{1-\varepsilon^2}\leqslant 4\exp\left(2\varepsilon^2\right)$, *where the inequality holds provided that* $|\varepsilon|\leqslant 1/2$.

*Proof.* Starting with (a), note that

$$\begin{aligned}\langle\mathbf{x},\mathbf{y}\rangle &= \sum_{i\in[n]\setminus\mathrm{Diff}(\mathbf{x},\mathbf{y})}\mathbf{x}_i\mathbf{y}_i+\sum_{i\in\mathrm{Diff}(\mathbf{x},\mathbf{y})}\mathbf{x}_i\mathbf{y}_i\\ &= (n-|\mathrm{Diff}(\mathbf{x},\mathbf{y})|)-|\mathrm{Diff}(\mathbf{x},\mathbf{y})|\\ &= n-2|\mathrm{Diff}(\mathbf{x},\mathbf{y})|.\end{aligned}$$

It thus follows that $\frac{|\mathrm{Diff}(\mathbf{x},\mathbf{y})|}{n}=\frac{n-\langle\mathbf{x},\mathbf{y}\rangle}{2n}=\frac{1-\varepsilon}{2}$. Hence

$$\alpha=1-\frac{|\mathrm{Diff}(\mathbf{x},\mathbf{y})|}{n}=1-\frac{1-\varepsilon}{2}=\frac{1+\varepsilon}{2}.$$

Next, we prove (b). Using (a) we obtain

$$\frac{1}{\alpha(1-\alpha)}=\frac{1}{\frac{1+\varepsilon}{2}\cdot\frac{1-\varepsilon}{2}}=\frac{4}{(1+\varepsilon)(1-\varepsilon)}=\frac{4}{1-\varepsilon^2}.$$

Assume now that $|\varepsilon| \leqslant 1/2$. Since $1 - x \geqslant \exp(-2x)$ holds whenever $0 \leqslant x \leqslant 1/2$, it follows that

$$\frac{4}{1-\varepsilon^2} \leqslant \frac{4}{\exp\left(-2\varepsilon^2\right)} = 4\exp\left(2\varepsilon^2\right). \qquad \square$$

A key tool in our approach is the following approximation result for binomial coefficients $\binom{n}{k}$, where $k$ is "close" to $n/2$.

**Proposition 11.** *Let $n$ be a sufficiently large even integer and let $t \in \mathbb{Z}$ be such that $|t| = o(n)$ and $\frac{n+t}{2} \in \mathbb{Z}$. Then,*

$$\binom{n}{\frac{n+t}{2}} = (1+o_n(1))\sqrt{\frac{2}{\pi n}} \cdot 2^n \exp\left(-\frac{t^2}{2n} + \Theta\left(\frac{t^3}{n^2}\right) + o\left(\frac{t}{n}\right)\right). \tag{2}$$

*Remark* 12. Up to small modifications, Proposition 11 and its proof can be found in [46, Section 5.4]; we include the proposition and its proof here as these modifications are important for our purposes.

*Proof of Proposition 11.* Let

$$Q = \binom{n}{\frac{n+t}{2}} \Big/ \binom{n}{n/2} = \frac{(n/2)!(n/2)!}{\left(\frac{n+t}{2}\right)!\left(\frac{n-t}{2}\right)!} = \prod_{j=1}^{t/2} \frac{n/2-j+1}{n/2+j}.$$

Therefore

$$\log Q = \sum_{j=1}^{t/2} \log\left(1 - \frac{4j-2}{n+2j}\right) = \sum_{j=1}^{t/2} \left[-\frac{4j-2}{n+2j} + \Theta\left(\frac{j^2}{n^2}\right)\right], \tag{3}$$

where for the last equality we use the expansion $\log(1-x) = -x + \Theta(x^2)$, holding whenever $x \in (0,1)$. Substituting the identity

$$\frac{4j-2}{n+2j} = \frac{4j}{n} - \frac{8j^2}{n(n+2j)} - \frac{2}{n+2j} = \frac{4j}{n} - \frac{2}{n+2j} + \Theta\left(\frac{j^2}{n^2}\right)$$

into (3) yields

$$\log Q = -\sum_{j=1}^{t/2} \frac{4j}{n} + \sum_{j=1}^{t/2} \frac{2}{n+2j} + \sum_{j=1}^{t/2} \Theta\left(\frac{j^2}{n^2}\right) = -\frac{t}{n} - \frac{t^2}{2n} + \sum_{j=1}^{t/2} \frac{2}{n+2j} + \Theta(t^3/n^2), \tag{4}$$

where for the last equality we employ the identity $\sum_{i=1}^{k} i = k(k+1)/2$ and the estimate $\sum_{i=1}^{k} i^2 = \Theta(k^3)$.

The sum appearing on the right hand side of (4) satisfies

$$\frac{t}{n+t} = \sum_{j=1}^{t/2} \frac{2}{n+t} \leqslant \sum_{j=1}^{t/2} \frac{2}{n+2j} \leqslant \sum_{j=1}^{t/2} \frac{2}{n} = \frac{t}{n}.$$

Since $t = o(n)$, it follows that

$$\sum_{j=1}^{t/2} \frac{2}{n+2j} = (1+o(1))t/n = t/n + o(t/n). \tag{5}$$

Combining (4) and (5) then implies that

$$\log Q = -\frac{t^2}{2n} + \Theta\left(\frac{t^3}{n^2}\right) + o\left(\frac{t}{n}\right).$$

The claim follows since

$$\binom{n}{n/2} = (1+o_n(1))\sqrt{\frac{2}{\pi n}} \cdot 2^n$$

holds by a straightforward application of Stirling's approximation[10]. □

The following lemma captures yet another crucial difference between our argument fitting the discrete scenario and that seen for the Gaussian case in [16]. Lemma 13 associates with each member $\mathbf{x}$ of the support of a certain deterministic extension $\mathcal{S}$ of $\mathcal{D}$, a vector $\mathbf{s}^{\mathbf{x}}$ capturing the point-wise gaps between $M\mathbf{x}$ and $R\mathbf{x}$; see (9) for details.

**Lemma 13.** *Let $i \in \{1,2\}$ and let $M \in \mathbb{R}^{d\times n}$ be a Komlós matrix whose last $i$ columns are $\mathbf{0}$. Let $\mathcal{S}$ be a distribution over $\{-1,1\}^n$ which forms a deterministic extension of $\mathcal{D}_{n-i}$. Then, for every $\mathbf{x} \in \operatorname{Supp}\mathcal{S}$, there exists a vector $\mathbf{s}^{\mathbf{x}} = (\mathbf{s}_1^{\mathbf{x}}, \dots \mathbf{s}_d^{\mathbf{x}})$ which satisfies the following three properties.*

(1) $\mathbf{s}_i^{\mathbf{x}} \in [-4,4]$ *for every $i \in [d]$;*

(2) $-\sqrt{d}(M\mathbf{x})_i + \mathbf{s}_i^{\mathbf{x}} \equiv n \pmod 4$ *for every $i \in [d]$;*

(3) $\left|\sum_{i=1}^d \mathbf{s}_i^{\mathbf{x}}(M\mathbf{x})_i\right| = O(\sqrt{\log d})$.

*Proof.* Note first that for every positive integer $n$ and any real number $a$ there exists a unique real number $a' \in [0,4)$ such that $a + a' \equiv n \pmod 4$; let $f_n : \mathbb{R} \to \mathbb{R}$ be defined by $f_n(a) = a'$. Given any $\mathbf{x} \in \mathcal{S}$ we determine the coordinates of $\mathbf{s}^{\mathbf{x}}$ sequentially. Set $\mathbf{s}_1^{\mathbf{x}} = f_n(-\sqrt{d}(M\mathbf{x})_1)$. Suppose we have already determined $\mathbf{s}_1^{\mathbf{x}}, \dots \mathbf{s}_j^{\mathbf{x}}$ for some $j \in [d-1]$ and now aim to choose $\mathbf{s}_{j+1}^{\mathbf{x}}$. If $(M\mathbf{x})_{j+1} \cdot \sum_{i=1}^j \mathbf{s}_i^{\mathbf{x}}(M\mathbf{x})_i < 0$, then we set $\mathbf{s}_{j+1}^{\mathbf{x}} = f_n(-\sqrt{d}(M\mathbf{x})_{j+1})$; in all other cases we set $\mathbf{s}_{j+1}^{\mathbf{x}} = f_n(-\sqrt{d}(M\mathbf{x})_{j+1}) - 4$. Observe that Properties (1) and (2) follow immediately from the definition of $f_n$ and from our process of choosing $\mathbf{s}_1^{\mathbf{x}}, \dots \mathbf{s}_d^{\mathbf{x}}$. Similarly, Property (3) follows from our process of choosing $\mathbf{s}_1^{\mathbf{x}}, \dots \mathbf{s}_d^{\mathbf{x}}$ and by Lemma 9(ii). □

[10] Use $\sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{1/(12n+1)} \leqslant n! < \sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{1/12n}$.

### 2.2 Rudimentary probabilistic estimations

The main results of this section are Lemmas 15 and 16 stated below. The proofs of these being rudimentary (yet crucial to subsequent arguments) are thus delegated to Appendix A. Roughly put, these two lemmas deal with determining the probabilities of events of the form $\langle \mathbf{r}, \mathbf{x} \rangle = 2k$, where $\mathbf{r}$ is a Rademacher vector, $\mathbf{x} \in \{-1,1\}^n$, $n \in \mathbb{N}$ is even, and $k \in \mathbb{Z}$; we refer to such probabilities as *core probabilities*. The focus on the inner product being even is owing to the fact that $\sum_{i=1}^{n} \mathbf{y}_i = \#_1(\mathbf{y}) - \#_{-1}(\mathbf{y})$ holds for any vector $\mathbf{y} \in \{-1,1\}^n$. Since $n$ is even, there exists an integer $y$ such that $\#_1(\mathbf{y}) = n/2 + y$ leading to $\sum_{i=1}^{n} \mathbf{y}_i = n/2 + y - (n/2 - y) = 2y$. The following is then implied.

**Observation 14.** *Let $n$ be a positive even integer and let $k \in \mathbb{Z}$. Then,*

$$|S_k| = \binom{n}{n/2+k}, \tag{6}$$

*where* $S_k := \left\{ \mathbf{v} \in \{-1,1\}^n : \sum_{i=1}^{n} \mathbf{v}_i = 2k \right\}$.

Let

$$\mathcal{E}_n = \left\{ \mathbf{v} \in \{-1,1\}^n : \#_1(\mathbf{v}) \equiv 0 \pmod 2 \right\}$$

denote the set of so-called *even* members of $\{-1,1\}^n$. The first main result of this section reads as follows.

**Lemma 15.** *Let $n \in \mathbb{N}$ be even, let $\mathbf{r}$ be a vector sampled uniformly at random from $\mathcal{E}_n$, let $\mathbf{x} \in \{-1,1\}^n$, and let $k \in \mathbb{Z}$ be such that $\mathbb{P}[\langle \mathbf{r}, \mathbf{x} \rangle = 2k] > 0$. Then,*

$$\mathbb{P}\Big[\langle \mathbf{r}, \mathbf{x} \rangle = 2k\Big] = \frac{1}{2^{n-1}} \binom{n}{n/2+k}. \tag{7}$$

The second main result of this section reads as follows.

**Lemma 16.** *Let $n \in \mathbb{N}$ be even, let $\mathbf{r}$ be a vector sampled uniformly at random from $\mathcal{E}_n$, let $\mathbf{x}, \mathbf{y} \in \{-1,1\}^n$ satisfying $\#_1(\mathbf{x}) \equiv \#_1(\mathbf{y}) \pmod 2$ be given, and let $\alpha = \alpha(\mathbf{x}, \mathbf{y})$ be as in Claim 10. Then, for any pair of integers $k_{\mathbf{x}}$ and $k_{\mathbf{y}}$ satisfying $\mathbb{P}\left[\langle \mathbf{r}, \mathbf{x} \rangle = 2k_{\mathbf{x}}, \langle \mathbf{r}, \mathbf{y} \rangle = 2k_{\mathbf{y}}\right] > 0$, the equality*

$$\mathbb{P}\Big[\langle \mathbf{r}, \mathbf{x} \rangle = 2k_{\mathbf{x}}, \langle \mathbf{r}, \mathbf{y} \rangle = 2k_{\mathbf{y}}\Big] = \frac{1}{2^{n-1}} \binom{\alpha n}{\frac{\alpha n + k_{\mathbf{x}} + k_{\mathbf{y}}}{2}} \binom{(1-\alpha)n}{\frac{(1-\alpha)n + k_{\mathbf{x}} - k_{\mathbf{y}}}{2}} \tag{8}$$

*holds.*

## 3 Proof of the main result - Theorem 1

Using the fact that $\|M\|_\infty \leqslant 1$ holds whenever $M$ is Komlós (for indeed $\|\mathbf{v}\|_\infty \leqslant \|\mathbf{v}\|_2 \leqslant 1$ holds for every column $\mathbf{v}$ of $M$), we deduce Theorem 1 from the following claim.

**Claim 17.** *Let $d = \omega(1)$ be an integer and let $n = \omega(d \log d)$ be an even integer. Let $\mathcal{S}$ be a distribution over $\{-1,1\}^n$ which forms a deterministic extension of $\mathcal{D}_{n-i}$, for some $i \in \{1,2\}$, and such that $\#_1(\mathbf{x}) \equiv 0 \pmod 2$ holds for every vector $\mathbf{x} \in \operatorname{Supp} \mathcal{S}$. Let $M \in \mathbb{R}^{d\times n}$ be a Komlós matrix whose last $i$ columns are $\mathbf{0}$, and let $R \in \mathbb{R}^{d\times n}$ be a Rademacher matrix, conditioned on $\#_1(\mathbf{r}) \equiv 0 \pmod 2$ holding for every row $\mathbf{r}$ of $R$. Then, a.a.s. there exists a vector $\mathbf{x} \in \operatorname{Supp} \mathcal{S}$ such that $\|(M + R/\sqrt{d})\mathbf{x}\|_\infty \leqslant 4d^{-1/2}$ holds.*

**Claim 17 implies Theorem 1:** Let $n$ and $M$ per the premise of Theorem 1 be given. Set $M_1 := [\ M \mid \mathbf{0}\ ] \in \mathbb{R}^{d\times(n+1)}$ and $M_2 := [\ M \mid \mathbf{0} \mid \mathbf{0}\ ] \in \mathbb{R}^{d\times(n+2)}$, where $\mathbf{0}$ denotes the zero vector in $\mathbb{R}^d$; in particular, $M_1$ and $M_2$ are both Komlós. Let $R_1 \in \mathbb{R}^{d\times(n+1)}$ and $R_2 \in \mathbb{R}^{d\times(n+2)}$ be Rademacher matrices, each satisfying the row parity condition stated in Claim 17.

Given $\mathbf{x} \in \operatorname{Supp} \mathcal{D}$, define $\mathbf{x}^{(1)} := [\mathbf{x} \mid \ell] \in \{-1,1\}^{n+1}$ and $\mathbf{x}^{(2)} := [\mathbf{x} \mid \ell_1 \mid \ell_2] \in \{-1,1\}^{n+2}$, where

$$\ell := \begin{cases} -1, & \#_1(\mathbf{x}) \equiv 0 \pmod 2, \\ 1, & \#_1(\mathbf{x}) \equiv 1 \pmod 2, \end{cases}$$

and

$$(\ell_1, \ell_2) := \begin{cases} (-1,-1), & \#_1(\mathbf{x}) \equiv 0 \pmod 2, \\ (-1,\ 1), & \#_1(\mathbf{x}) \equiv 1 \pmod 2. \end{cases}$$

It follows that $\#_1(\mathbf{x}^{(1)}) \equiv \#_1(\mathbf{x}^{(2)}) \equiv 0 \pmod 2$ holds for every $\mathbf{x} \in \operatorname{Supp} \mathcal{D}$. For $i \in \{1,2\}$, define $\mathcal{S}_i$ to be a distribution set over $\{-1,1\}^{n+i}$ obtained by first sampling a vector $\mathbf{x} \in \{-1,1\}^n$ according to the distribution $\mathcal{D}$ and then performing the (injective) deterministic extension yielding $\mathbf{x}^{(i)}$.

If $n$ is odd, then set $N := M_1$, $\mathcal{S} = \mathcal{S}_1$, and $R := R_1$; otherwise set $N := M_2$, $\mathcal{S} = \mathcal{S}_2$, and $R = R_2$. Claim 17 asserts that a.a.s. there exists a vector $\mathbf{y} \in \operatorname{Supp} \mathcal{S}$ for which $\|(N + R/\sqrt{d})\mathbf{y}\|_\infty \leqslant 4d^{-1/2}$ holds. Resampling the first entry of every row of $R$ allows for a conformal Rademacher matrix to be sampled uniformly at random at the price of increasing the discrepancy by at most $2d^{-1/2}$ asymptotically almost surely. Expose $R$ and let $R'$ be the matrix obtained from $R$ by dropping its last column, if $n$ is odd, and its last two columns, if $n$ is even. In addition, let $\mathbf{y}' \in \{-1,1\}^n$ be the vector obtained from $\mathbf{y}$ by dropping its last entry, if $n$ is odd, and its last two entries, if $n$ is even. Note that, $\|(M + R'/\sqrt{d})\mathbf{y}'\|_\infty \leqslant 8d^{-1/2}$. □

The remainder of this section is devoted to the proof of Claim 17. For every $\mathbf{x} \in \operatorname{Supp} \mathcal{S}$ and every $i \in [d]$, let $\mathbf{s}_i^{\mathbf{x}}$ be as in Lemma 13 and let $\mathbf{t}_i^{\mathbf{x}} = -\sqrt{d}(M\mathbf{x})_i + \mathbf{s}_i^{\mathbf{x}}$. Set $\Delta := 4d^{-1/2}$ and define the random variable

$$S := S(R) = \sum_{\mathbf{x}\in\operatorname{Supp}\mathcal{S}} \mathbb{1}\left\{R\mathbf{x} = (\mathbf{t}_1^{\mathbf{x}}, \ldots, \mathbf{t}_d^{\mathbf{x}})\right\} \cdot \mathbb{P}_{\mathbf{y}\sim\mathcal{S}}[\mathbf{y} = \mathbf{x}] = \mathbb{E}_{\mathbf{x}\sim\mathcal{S}}\left[\mathbb{1}\left\{R\mathbf{x} = (\mathbf{t}_1^{\mathbf{x}}, \ldots, \mathbf{t}_d^{\mathbf{x}})\right\}\right] \tag{9}$$

whose sole source of randomness is $R$. It suffices to prove that $S > 0$ holds asymptotically almost surely. Indeed, if the latter holds, then for *almost every* Rademacher matrix $R$, there exists a vector $\mathbf{x} \in \operatorname{Supp} \mathcal{S}$ for which

$$\mathbb{1}\left\{R\mathbf{x} = (\mathbf{t}_1^{\mathbf{x}}, \ldots, \mathbf{t}_d^{\mathbf{x}})\right\} \cdot \mathbb{P}_{\mathbf{y}\sim\mathcal{S}}[\mathbf{y} = \mathbf{x}] > 0$$

holds. It then follows by our choice of $\Delta$ and by Lemma 13(1) that for almost every Rademacher matrix $R$, there exists a vector $\mathbf{x} \in \operatorname{Supp} \mathcal{S}$ for which the event $\left\|\left(M + R/\sqrt{d}\right)\mathbf{x}\right\|_\infty \leqslant \Delta$ occurs.

Establishing that $\mathbb{E}_R[S] > 0$ (in Claim 19 below) enables an appeal to the following consequence of the Paley-Zygmund inequality (see, e.g., [30])

$$\mathbb{P}_R[S > 0] \geqslant \frac{\mathbb{E}_R[S]^2}{\mathbb{E}_R[S^2]}. \tag{10}$$

Hence, given that $\mathbb{E}_R[S] > 0$ holds, it suffices to prove that

$$\mathbb{E}_R[S^2] \leqslant (1 + o(1))\,\mathbb{E}_R[S]^2 \tag{11}$$

in order to deduce that $\mathbb{P}_R[S > 0] \geqslant 1 - o(1)$.

Prior to proving Claim 19, it will be useful to establish the following simple fact.

**Claim 18.** $\mathbb{P}_R\left[R\mathbf{x} = (\mathbf{t}_1^{\mathbf{x}}, \ldots, \mathbf{t}_d^{\mathbf{x}})\right] > 0$ *for every* $\mathbf{x} \in \operatorname{Supp} \mathcal{S}$.

*Proof.* Fix an arbitrary vector $\mathbf{x} \in \operatorname{Supp} \mathcal{S}$. Then,

$$\|M\mathbf{x}\|_\infty = O(\sqrt{\log d}) < n/\sqrt{d},$$

where the equality holds by Lemma 9(ii), and the inequality holds since $n$ is assumed to be sufficiently large with respect to $d$. It follows that $(M\mathbf{x})_i \in [-n/\sqrt{d}, n/\sqrt{d}]$ holds for every $i \in [d]$.

Since $n$ is even and $\#_1(\mathbf{r}) \equiv 0 \pmod 2$ holds for every row $\mathbf{r}$ of $R$, it follows that for every $i \in [d]$ and every $k \in \{m \in [-n, n] : m \equiv n \pmod 4\}$, there exists a vector $\mathbf{r}_i \in \mathcal{E}_n$ such that $\langle \mathbf{r}_i, \mathbf{x}\rangle = k$. It then follows by Lemma 13(2) that there exists a choice of $R$ with each of its rows satisfying the parity condition stated in Claim 17 such that $(R\mathbf{x})_i = \mathbf{t}_i^{\mathbf{x}}$ holds for every $i \in [d]$; this concludes the proof of the claim. □

**Claim 19.** $\mathbb{E}_R[S] > 0$.

*Proof.* Note that

$$\mathbb{E}_R[S] = \mathbb{E}_{\mathbf{x}\sim\mathcal{S}}\,\mathbb{E}_R\left[\mathbb{1}\left\{R\mathbf{x} = (\mathbf{t}_1^{\mathbf{x}}, \ldots, \mathbf{t}_d^{\mathbf{x}})\right\}\right] = \mathbb{E}_{\mathbf{x}\sim\mathcal{S}}\,\mathbb{P}_R\left[R\mathbf{x} = (\mathbf{t}_1^{\mathbf{x}}, \ldots, \mathbf{t}_d^{\mathbf{x}})\right] > 0,$$

where the above inequality holds by Claim 18. □

Turning our attention to (11), note that

$$(\mathbb{E}_R[S])^2 = (\mathbb{E}_{\mathbf{x}\sim\mathcal{S}}\,\mathbb{P}_R\left[R\mathbf{x} = (\mathbf{t}_1^{\mathbf{x}},\dots,\mathbf{t}_d^{\mathbf{x}})\right]) \cdot (\mathbb{E}_{\mathbf{y}\sim\mathcal{S}}\,\mathbb{P}_R\left[R\mathbf{y} = (\mathbf{t}_1^{\mathbf{y}},\dots,\mathbf{t}_d^{\mathbf{y}})\right]) = \mathbb{E}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}\left[P_{\mathbf{x}}P_{\mathbf{y}}\right],$$

where, for every $\mathbf{x} \in \operatorname{Supp}\mathcal{S}$,

$$P_{\mathbf{x}} := \mathbb{P}_R\left[R\mathbf{x} = (\mathbf{t}_1^{\mathbf{x}},\dots,\mathbf{t}_d^{\mathbf{x}})\right].$$

Similarly

$$\begin{aligned}\mathbb{E}_R[S^2] &= \mathbb{E}_R\left[\mathbb{E}_{\mathbf{x}\sim\mathcal{S}}\left[\mathbb{1}\left\{R\mathbf{x} = (\mathbf{t}_1^{\mathbf{x}},\dots,\mathbf{t}_d^{\mathbf{x}})\right\}\right] \cdot \mathbb{E}_{\mathbf{y}\sim\mathcal{S}}\left[\mathbb{1}\left\{R\mathbf{y} = (\mathbf{t}_1^{\mathbf{y}},\dots,\mathbf{t}_d^{\mathbf{y}})\right\}\right]\right] \\ &= \mathbb{E}_R\,\mathbb{E}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}\left[\mathbb{1}\left\{R\mathbf{x} = (\mathbf{t}_1^{\mathbf{x}},\dots,\mathbf{t}_d^{\mathbf{x}})\right\} \cdot \mathbb{1}\left\{R\mathbf{y} = (\mathbf{t}_1^{\mathbf{y}},\dots,\mathbf{t}_d^{\mathbf{y}})\right\}\right] \\ &= \mathbb{E}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}\left[\mathbb{P}_R\left[R\mathbf{x} = (\mathbf{t}_1^{\mathbf{x}},\dots,\mathbf{t}_d^{\mathbf{x}}), R\mathbf{y} = (\mathbf{t}_1^{\mathbf{y}},\dots,\mathbf{t}_d^{\mathbf{y}})\right]\right] = \mathbb{E}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}\left[P_{\mathbf{x},\mathbf{y}}\right],\end{aligned}$$

where, for every $\mathbf{x},\mathbf{y} \in \operatorname{Supp}\mathcal{S}$,

$$P_{\mathbf{x},\mathbf{y}} := \mathbb{P}_R\left[R\mathbf{x} = (\mathbf{t}_1^{\mathbf{x}},\dots,\mathbf{t}_d^{\mathbf{x}}), R\mathbf{y} = (\mathbf{t}_1^{\mathbf{y}},\dots,\mathbf{t}_d^{\mathbf{y}})\right].$$

The goal (11) can then be rewritten as follows

$$\mathbb{E}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}\left[P_{\mathbf{x},\mathbf{y}}\right] \leqslant (1+o(1))\,\mathbb{E}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}\left[P_{\mathbf{x}}P_{\mathbf{y}}\right]. \tag{12}$$

We begin by considering the right hand side of (12). Our first result in this respect is an estimation of $P_{\mathbf{x}}$.

**Lemma 20.** *Suppose that $n = \omega(d)$. Then, for every $\mathbf{x} \in \operatorname{Supp}\mathcal{S}$, it holds that*

$$P_{\mathbf{x}} = (1+o_d(1))\left(\frac{8}{\pi n}\right)^{d/2} \cdot \exp\left(-\frac{1}{2n}\sum_{i=1}^{d}(\mathbf{t}_i^{\mathbf{x}})^2\right) \cdot \exp(\delta_{\mathbf{x}}),$$

*where* $\delta_{\mathbf{x}} = \left(O\left(n^{-2}d^{3/2}\log d\right) + o\left(n^{-1}\sqrt{d}\right)\right)|\langle M\mathbf{x},\mathbf{1}\rangle|$.

*Proof.* Fix any $\mathbf{x} \in \operatorname{Supp}\mathcal{S}$. It follows by the independence of the entries of $R$ that

$$\begin{aligned}P_{\mathbf{x}} &= \mathbb{P}_R\left[R\mathbf{x} = (\mathbf{t}_1^{\mathbf{x}},\dots,\mathbf{t}_d^{\mathbf{x}})\right] = \prod_{i=1}^{d}\mathbb{P}_R\left[(R\mathbf{x})_i = t_i^{\mathbf{x}}\right] = \prod_{i=1}^{d}\frac{1}{2^{n-1}}\binom{n}{\frac{n+\mathbf{t}_i^{\mathbf{x}}}{2}} \\ &= (1+o_d(1))\left(\frac{8}{\pi n}\right)^{d/2}\prod_{i=1}^{d}\exp\left(\Theta\left(\frac{(\mathbf{t}_i^{\mathbf{x}})^3}{n^2}\right)\right)\exp\left(o\left(\frac{\mathbf{t}_i^{\mathbf{x}}}{n}\right)\right)\exp\left(-\frac{(\mathbf{t}_i^{\mathbf{x}})^2}{2n}\right), \qquad (13)\end{aligned}$$

where the penultimate equality holds by Lemma 15 and the last equality holds by Proposition 11 (note that $\mathbf{t}_i^{\mathbf{x}} = o(n)$ holds by Lemma 9(ii) and that $(1+o_n(1))^d = 1+o_d(1)$ holds by footnote 8 and since $n = \omega(d\log d)$).

In light of (13), in order to complete the proof of the lemma, it suffices to prove that $\left|\sum_{i=1}^{d}\mathbf{t}_i^{\mathbf{x}}\right| = \sqrt{d}|\langle M\mathbf{x},\mathbf{1}\rangle| + o(n)$ and that $\left|\sum_{i=1}^{d}(\mathbf{t}_i^{\mathbf{x}})^3\right| = O\left(d^{3/2}\log d|\langle M\mathbf{x},\mathbf{1}\rangle|\right) + o(n^2)$, which would in turn imply that

$$\begin{aligned}\prod_{i=1}^{d}\exp\left(\Theta\left(\frac{(\mathbf{t}_i^{\mathbf{x}})^3}{n^2}\right)\right)\exp\left(o\left(\frac{\mathbf{t}_i^{\mathbf{x}}}{n}\right)\right) &= \exp\left(\Theta\left(\frac{\sum_{i=1}^{d}(\mathbf{t}_i^{\mathbf{x}})^3}{n^2}\right)+o\left(\frac{\sum_{i=1}^{d}\mathbf{t}_i^{\mathbf{x}}}{n}\right)\right)\\ &= (1+o_d(1))\exp\left(\left(O\left(n^{-2}d^{3/2}\log d\right)+o\left(n^{-1}\sqrt{d}\right)\right)|\langle M\mathbf{x},\mathbf{1}\rangle|\right).\end{aligned}$$

Since $\mathbf{t}_i^{\mathbf{x}} = -\sqrt{d}(M\mathbf{x})_i + \mathbf{s}_i^{\mathbf{x}}$ holds for every $i\in[d]$, it follows that

$$\left|\sum_{i=1}^{d}\mathbf{t}_i^{\mathbf{x}}\right| = \left|\sum_{i=1}^{d}-\sqrt{d}(M\mathbf{x})_i+\mathbf{s}_i^{\mathbf{x}}\right| \leqslant \sqrt{d}\left|\sum_{i=1}^{d}(M\mathbf{x})_i\right| + O(d) = \sqrt{d}|\langle M\mathbf{x},\mathbf{1}\rangle| + o(n), \tag{14}$$

where the last equality holds since $n=\omega(d)$ by the premise of the lemma.

Similarly, for every $i\in[d]$,

$$(\mathbf{t}_i^{\mathbf{x}})^3 = -d^{3/2}((M\mathbf{x})_i)^3 + 3d\mathbf{s}_i^{\mathbf{x}}((M\mathbf{x})_i)^2 - 3\sqrt{d}(\mathbf{s}_i^{\mathbf{x}})^2(M\mathbf{x})_i + (\mathbf{s}_i^{\mathbf{x}})^3.$$

It thus follows that

$$\begin{aligned}\left|\sum_{i=1}^{d}(\mathbf{t}_i^{\mathbf{x}})^3\right| &\leqslant d^{3/2}\left|\sum_{i=1}^{d}((M\mathbf{x})_i)^3\right| + 3d\left|\sum_{i=1}^{d}\mathbf{s}_i^{\mathbf{x}}((M\mathbf{x})_i)^2\right| - 3\sqrt{d}\left|\sum_{i=1}^{d}(\mathbf{s}_i^{\mathbf{x}})^2(M\mathbf{x})_i\right| + \left|\sum_{i=1}^{d}(\mathbf{s}_i^{\mathbf{x}})^3\right|\\ &\leqslant d^{3/2}\|M\mathbf{x}\|_\infty^2\left|\sum_{i=1}^{d}(M\mathbf{x})_i\right| + 12d\|M\mathbf{x}\|_2^2 + 48\sqrt{d}\left|\sum_{i=1}^{d}((M\mathbf{x})_i)^2+1\right| + O(d)\\ &\leqslant d^{3/2}\log d|\langle M\mathbf{x},\mathbf{1}\rangle| + O(d^2) + 48\sqrt{d}\|M\mathbf{x}\|_2^2 + O(d^{3/2}) + O(d)\\ &\leqslant d^{3/2}\log d|\langle M\mathbf{x},\mathbf{1}\rangle| + o(n^2),\end{aligned} \tag{15}$$

where the second inequality holds since $|\mathbf{s}_i^{\mathbf{x}}|\leqslant 4$ by Lemma 13(1), the penultimate inequality holds by parts (i) and (ii) of Lemma 9, and the last inequality holds by Lemma 9(i) and since $n=\omega(d)$ by the premise of the lemma. □

Lemma 20 provides the following useful uniform estimation on the probabilities $P_{\mathbf{x}}$.

**Lemma 21.** *Suppose that $n=\omega(d)$ and let $p = \left(\frac{8}{\pi n}\right)^{d/2}\cdot\exp\left(-r^2\delta^2/2\right)$, where $r$ is as in Lemma 7(i) and $\delta=\sqrt{d/n}$. Then, $P_{\mathbf{x}} = (1+o_d(1))p\cdot\exp\left(\delta_{\mathbf{x}}\right)$ holds for every $\mathbf{x}\in\operatorname{Supp}\mathcal{S}$, where $\delta_{\mathbf{x}}$ is as in Lemma 20.*

*Remark* 22. We refer above to the estimation of $P_{\mathbf{x}}$, appearing in Lemma 21, as being uniform even though it does depend on $\mathbf{x}$ due to the term $\exp\left(\delta_{\mathbf{x}}\right)$. This is because (the asymptotic value of the expectation of) this term turns out to be essentially 1 (see Lemma 25). An analogous result for the Gaussian case (with a somewhat simpler expression for $\delta_{\mathbf{x}}$) appears in [16] (see Claim 2.3 there).

*Proof of Lemma 21.* Fix any $\mathbf{x} \in \operatorname{Supp} \mathcal{S}$. In light of Lemma 20 it suffices to show that

$$\frac{1}{2n}\sum_{i=1}^{d}(t_i^{\mathbf{x}})^2 = r^2\delta^2/2 + o_d(1).$$

Since $t_i^{\mathbf{x}} = -\sqrt{d}(M\mathbf{x})_i + s_i^{\mathbf{x}}$, where $s_i^{\mathbf{x}} \in [-4,4]$ by Lemma 13(1), holds for every $i \in [d]$, it follows that

$$\begin{aligned}\sum_{i=1}^{d}(t_i^{\mathbf{x}})^2 &= d\sum_{i=1}^{d}((M\mathbf{x})_i)^2 - 2\sqrt{d}\sum_{i=1}^{d}s_i^{\mathbf{x}}(M\mathbf{x})_i + \sum_{i=1}^{d}(s_i^{\mathbf{x}})^2 \\ &= d\|M\mathbf{x}\|_2^2 - 2\sqrt{d}\sum_{i=1}^{d}s_i^{\mathbf{x}}(M\mathbf{x})_i + O(d), \end{aligned} \tag{16}$$

where the last equality holds by Lemma 13(1). For $r = O(\sqrt{d})$ and $\mathcal{T} = d^{-C'}$, with $C' > 1$ some constant, $\|M\mathbf{x}\|_2 \in [r-\mathcal{T}, r+\mathcal{T}]$ holds, by Lemma 9(i); this, in turn, implies that

$$r^2 - O(1) \leqslant (r-\mathcal{T})^2 \leqslant \|M\mathbf{x}\|_2^2 \leqslant (r+\mathcal{T})^2 \leqslant r^2 + O(1).$$

Since, moreover, $n = \omega(d)$, $r = O(\sqrt{d})$, and $\delta = \sqrt{d/n}$, it follows that

$$\delta^2\|M\mathbf{x}\|_2^2/2 = \delta^2 r^2/2 \pm O\left(dn^{-1}\right) = \delta^2 r^2/2 + o_d(1). \tag{17}$$

Combining (16) and (17), leads to

$$\begin{aligned}\frac{1}{2n}\sum_{i=1}^{d}(t_i^{\mathbf{x}})^2 &= \frac{1}{2n}d\|M\mathbf{x}\|_2^2 - \frac{1}{n}\sqrt{d}\sum_{i=1}^{d}s_i^{\mathbf{x}}(M\mathbf{x})_i + O\left(n^{-1}d\right) \\ &= \delta^2 r^2/2 - \frac{1}{n}\sqrt{d}\sum_{i=1}^{d}s_i^{\mathbf{x}}(M\mathbf{x})_i + o_d(1).\end{aligned}$$

In order to conclude the proof of the lemma, it remains to prove that $\left|n^{-1}\sqrt{d}\sum_{i=1}^{d}s_i^{\mathbf{x}}(M\mathbf{x})_i\right| = o_d(1)$. Indeed, it follows by Lemma 13(3) that

$$\left|n^{-1}\sqrt{d}\sum_{i=1}^{d}s_i^{\mathbf{x}}(M\mathbf{x})_i\right| = n^{-1}\sqrt{d}\left|\sum_{i=1}^{d}s_i^{\mathbf{x}}(M\mathbf{x})_i\right| = O\left(n^{-1}\sqrt{d\log d}\right) = o_d(1). \qquad \square$$

We turn our attention to the left hand side of (12).

**Lemma 23.** *Let $n = \omega(d)$ be an even integer and let $\delta = \sqrt{d/n}$. Let $\mathbf{x}, \mathbf{y} \in \operatorname{Supp}\mathcal{S}$ satisfying $-1/2 \leqslant \varepsilon := \varepsilon(\mathbf{x},\mathbf{y}) \leqslant 1/2$ be given, and let*

$$\begin{aligned}\beta(\mathbf{x},\mathbf{y}) := \exp&\left(d\varepsilon^2 + 2\delta^2|\varepsilon\langle M\mathbf{x}, M\mathbf{y}\rangle| + 2n^{-1}\sqrt{d}\,|\varepsilon\langle M\mathbf{x}, \mathbf{s}^{\mathbf{y}}\rangle| + 2n^{-1}\sqrt{d}\,|\varepsilon\langle M\mathbf{y}, \mathbf{s}^{\mathbf{x}}\rangle|\right) \\ &\cdot \exp\left(\left(O\left(n^{-2}d^{3/2}\log d\right) + o\left(n^{-1}\sqrt{d}\right)\right)\left(|\langle M\mathbf{x}, \mathbf{1}\rangle| + |\langle M\mathbf{y}, \mathbf{1}\rangle|\right)\right).\end{aligned}$$

*Then*

$$P_{\mathbf{x},\mathbf{y}} \leqslant (1+o_d(1))\cdot P_{\mathbf{x}}P_{\mathbf{y}}\cdot\beta(\mathbf{x},\mathbf{y})\cdot\exp\left(-\delta_{\mathbf{x}} - \delta_{\mathbf{y}}\right).$$

*Remark* 24. Ignoring the expression $\beta(\mathbf{x},\mathbf{y})\cdot\exp\left(-\delta_{\mathbf{x}}-\delta_{\mathbf{y}}\right)$, Lemma 23 provides an upper bound on $P_{\mathbf{x},\mathbf{y}}$ that constitutes a major step towards proving (12). This expression is handled formally by Lemmas 21 and 26. An analogous result for the Gaussian case (with a somewhat simpler expression for $\beta(\mathbf{x},\mathbf{y})$) appears in [16] (see Claim 2.2 there).

*Proof of Lemma 23.* Owing to our assumption that $|\varepsilon|\leqslant 1/2$, we may restrict our attention to pairs $(\mathbf{x},\mathbf{y})\in(\operatorname{Supp}\mathcal{S})^2$ such that $\mathbf{x}\neq\mathbf{y}$. Given such a pair, let $k_i^{\mathbf{x}}$ and $k_i^{\mathbf{y}}$ be integers satisfying $\mathbf{t}_i^{\mathbf{x}}=2k_i^{\mathbf{x}}$ and $\mathbf{t}_i^{\mathbf{y}}=2k_i^{\mathbf{y}}$. In a manner similar to that seen in the proof of Lemma 20, it holds that

$$\begin{aligned}P_{\mathbf{x},\mathbf{y}}&=\mathbb{P}_R\left[R\mathbf{x}=(\mathbf{t}_1^{\mathbf{x}},\ldots,\mathbf{t}_d^{\mathbf{x}}),R\mathbf{y}=(\mathbf{t}_1^{\mathbf{y}},\ldots,\mathbf{t}_d^{\mathbf{y}})\right]=\prod_{i=1}^{d}\mathbb{P}_R\left[(R\mathbf{x})_i=\mathbf{t}_i^{\mathbf{x}},\ (R\mathbf{y})_i=\mathbf{t}_i^{\mathbf{y}}\right]\\&=\prod_{i=1}^{d}\frac{1}{2^{n-1}}\binom{\alpha n}{\frac{\alpha n+k_i^{\mathbf{x}}+k_i^{\mathbf{y}}}{2}}\binom{(1-\alpha)n}{\frac{(1-\alpha)n+k_i^{\mathbf{x}}-k_i^{\mathbf{y}}}{2}},\end{aligned}\tag{18}$$

where the last equality holds by (8). For

$$L_i^{(1)}:=\binom{\alpha n}{\frac{\alpha n+k_i^{\mathbf{x}}+k_i^{\mathbf{y}}}{2}},$$

we obtain

$$L_i^{(1)}=\big(1+o_n(1)\big)2^{\alpha n}\sqrt{\frac{2}{\pi\alpha n}}\exp\left(\Theta\left(\frac{(k_i^{\mathbf{x}}+k_i^{\mathbf{y}})^3}{(\alpha n)^2}\right)+o\left(\frac{k_i^{\mathbf{x}}+k_i^{\mathbf{y}}}{\alpha n}\right)\right)\exp\left(-\frac{(k_i^{\mathbf{x}}+k_i^{\mathbf{y}})^2}{2\alpha n}\right),$$

where the equality holds by Proposition 11. Similarly, for

$$L_i^{(2)}:=\binom{(1-\alpha)n}{\frac{(1-\alpha)n+k_i^{\mathbf{x}}-k_i^{\mathbf{y}}}{2}},$$

we obtain

$$L_i^{(2)}=\big(1+o_n(1)\big)2^{(1-\alpha)n}\sqrt{\frac{2}{\pi(1-\alpha)n}}\exp\left(\Theta\left(\frac{(k_i^{\mathbf{x}}-k_i^{\mathbf{y}})^3}{((1-\alpha)n)^2}\right)+o\left(\frac{k_i^{\mathbf{x}}-k_i^{\mathbf{y}}}{(1-\alpha)n}\right)\right)\exp\left(-\frac{(k_i^{\mathbf{x}}-k_i^{\mathbf{y}})^2}{2(1-\alpha)n}\right).$$

It thus follows by (18) that

$$P_{\mathbf{x},\mathbf{y}}=\prod_{i=1}^{d}\frac{1}{2^{n-1}}L_i^{(1)}L_i^{(2)}=(1+o_d(1))\prod_{i=1}^{d}\frac{4}{\pi n}\sqrt{\frac{1}{\alpha(1-\alpha)}}\cdot\exp\left(D_i+E_i+F_i\right),\tag{19}$$

where

$$\begin{aligned}D_i&:=-\frac{(k_i^{\mathbf{x}}+k_i^{\mathbf{y}})^2}{2\alpha n}-\frac{(k_i^{\mathbf{x}}-k_i^{\mathbf{y}})^2}{2(1-\alpha)n};\\E_i&:=\Theta\left(\frac{(k_i^{\mathbf{x}}+k_i^{\mathbf{y}})^3}{(\alpha n)^2}+\frac{(k_i^{\mathbf{x}}-k_i^{\mathbf{y}})^3}{((1-\alpha)n)^2}\right);\end{aligned}$$

$$F_i := o\left(\frac{k_i^{\mathbf{x}}+k_i^{\mathbf{y}}}{\alpha n}+\frac{k_i^{\mathbf{x}}-k_i^{\mathbf{y}}}{(1-\alpha)n}\right).$$

In the sequel we prove that

$$\begin{aligned}\sum_{i=1}^{d}(D_i+E_i+F_i) \leqslant & -\frac{1}{2n}\left(\sum_{i=1}^{d}(\mathbf{t}_i^{\mathbf{x}})^2+\sum_{i=1}^{d}(\mathbf{t}_i^{\mathbf{y}})^2\right)+2\delta^2|\varepsilon\langle M\mathbf{x},M\mathbf{y}\rangle| \\ & +2n^{-1}\sqrt{d}\,|\varepsilon\langle M\mathbf{x},s^{\mathbf{y}}\rangle|+2n^{-1}\sqrt{d}\,|\varepsilon\langle M\mathbf{y},\mathbf{s}^{\mathbf{x}}\rangle| \\ & +O\left(n^{-2}d^{3/2}\log d\right)\left(|\langle M\mathbf{x},\mathbf{1}\rangle|+|\langle M\mathbf{y},\mathbf{1}\rangle|\right) \\ & +o\left(n^{-1}\sqrt{d}\right)\left(|\langle M\mathbf{x},\mathbf{1}\rangle|+|\langle M\mathbf{y},\mathbf{1}\rangle|\right)+o(1). \end{aligned} \tag{20}$$

Using (19) and (20) the proof concludes as follows. First, note that

$$\begin{aligned}P_{\mathbf{x},\mathbf{y}} &= (1+o_d(1))\prod_{i=1}^{d}\frac{4}{\pi n}\sqrt{\frac{1}{\alpha(1-\alpha)}}\cdot\exp\left(D_i+E_i+F_i\right) \\ &\leqslant (1+o_d(1))\left(\frac{8}{\pi n}\right)^d\exp\left(d\varepsilon^2+\sum_{i=1}^{d}(D_i+E_i+F_i)\right) \\ &\leqslant (1+o_d(1))\left(\frac{8}{\pi n}\right)^d\exp\left(-\frac{1}{2n}\left(\sum_{i=1}^{d}(\mathbf{t}_i^{\mathbf{x}})^2+\sum_{i=1}^{d}(\mathbf{t}_i^{\mathbf{y}})^2\right)\right)\cdot\beta(\mathbf{x},\mathbf{y}),\end{aligned}$$

where the first inequality holds by Claim 10(b).

Second, note that since $P_{\mathbf{x}}=(1+o_d(1))\left(\frac{8}{\pi n}\right)^{d/2}\cdot\exp\left(-\frac{1}{2n}\sum_{i=1}^{d}(\mathbf{t}_i^{\mathbf{x}})^2\right)\cdot\exp\left(\delta_{\mathbf{x}}\right)$ holds by Lemma 20, it follows that

$$P_{\mathbf{x},\mathbf{y}}\leqslant(1+o_d(1))P_{\mathbf{x}}P_{\mathbf{y}}\cdot\beta(\mathbf{x},\mathbf{y})\cdot\exp\left(-\delta_{\mathbf{x}}-\delta_{\mathbf{y}}\right),$$

as claimed. It remains to prove (20); to do so, we estimate each of the sums $\sum D_i$, $\sum E_i$, and $\sum F_i$ separately.

**Estimating $\sum \boldsymbol{D_i}$.** Start by writing

$$\begin{aligned}D_i &= -\frac{(1-\alpha)(k_i^{\mathbf{x}}+k_i^{\mathbf{y}})^2+\alpha(k_i^{\mathbf{x}}-k_i^{\mathbf{y}})^2}{2\alpha(1-\alpha)n} \\ &= -\frac{4(1-\alpha)\left((k_i^{\mathbf{x}})^2+2k_i^{\mathbf{x}}k_i^{\mathbf{y}}+(k_i^{\mathbf{y}})^2\right)+4\alpha\left((k_i^{\mathbf{x}})^2-2k_i^{\mathbf{x}}k_i^{\mathbf{y}}+(k_i^{\mathbf{y}})^2\right)}{2(1-\varepsilon^2)n} \\ &= -\frac{4\left((k_i^{\mathbf{x}})^2+(k_i^{\mathbf{y}})^2\right)}{2(1-\varepsilon^2)n}-\frac{4(1-2\alpha)\cdot 2k_i^{\mathbf{x}}k_i^{\mathbf{y}}}{2(1-\varepsilon^2)n} \\ &= -\frac{(\mathbf{t}_i^{\mathbf{x}})^2+(\mathbf{t}_i^{\mathbf{y}})^2}{2n}-\frac{\varepsilon^2\left((\mathbf{t}_i^{\mathbf{x}})^2+(\mathbf{t}_i^{\mathbf{y}})^2\right)}{2(1-\varepsilon^2)n}+\frac{\varepsilon\mathbf{t}_i^{\mathbf{x}}\mathbf{t}_i^{\mathbf{y}}}{(1-\varepsilon^2)n} \\ &\leqslant -\frac{(\mathbf{t}_i^{\mathbf{x}})^2+(\mathbf{t}_i^{\mathbf{y}})^2}{2n}+\frac{\varepsilon\mathbf{t}_i^{\mathbf{x}}\mathbf{t}_i^{\mathbf{y}}}{(1-\varepsilon^2)n},\end{aligned} \tag{21}$$

where the second equality holds by Claim 10(b); the last equality holds since $(1-2\alpha) = -\varepsilon$, $\mathbf{t}_i^{\mathbf{x}} = 2k_i^{\mathbf{x}}$ and $\mathbf{t}_i^{\mathbf{y}} = 2k_i^{\mathbf{y}}$, and by the following equality

$$\frac{A}{(1-z)B} = \frac{A}{B} + \frac{zA}{(1-z)B},$$

holding for every $A, B$ and $z$. Finally, the inequality follows by discarding the middle term appearing on the preceding line; the latter is negative owing to $(1-\varepsilon^2) \geqslant 1/2$, which holds since $|\varepsilon| \leqslant 1/2$.

For the term $\mathbf{t}_i^{\mathbf{x}}\mathbf{t}_i^{\mathbf{y}}$ appearing on the right hand side of (21), we may write

$$\begin{aligned}\mathbf{t}_i^{\mathbf{x}}\mathbf{t}_i^{\mathbf{y}} &= \left(-\sqrt{d}(M\mathbf{x})_i + \mathbf{s}_i^{\mathbf{x}}\right)\left(-\sqrt{d}(M\mathbf{y})_i + \mathbf{s}_i^{\mathbf{y}}\right)\\ &= d(M\mathbf{x})_i(M\mathbf{y})_i - \mathbf{s}_i^{\mathbf{y}}\sqrt{d}(M\mathbf{x})_i - \mathbf{s}_i^{\mathbf{x}}\sqrt{d}(M\mathbf{y})_i + \mathbf{s}_i^{\mathbf{x}}\mathbf{s}_i^{\mathbf{y}};\end{aligned}$$

indeed, for every $i \in [d]$ and $\mathbf{z} \in \{\mathbf{x}, \mathbf{y}\}$, the equality $\mathbf{t}_i^{\mathbf{z}} = -\sqrt{d}(M\mathbf{z})_i + \mathbf{s}_i^{\mathbf{z}}$ holds, where $\mathbf{s}_i^{\mathbf{z}} \in [-4, 4]$. Set

$$\mathcal{N} := \sum_{i=1}^{d} \frac{\sqrt{d}\varepsilon s_i^{\mathbf{y}}(M\mathbf{x})_i}{(1-\varepsilon^2)n} + \sum_{i=1}^{d} \frac{\sqrt{d}\varepsilon \mathbf{s}_i^{\mathbf{x}}(M\mathbf{y})_i}{(1-\varepsilon^2)n} - \sum_{i=1}^{d} \frac{\varepsilon \mathbf{s}_i^{\mathbf{x}}\mathbf{s}_i^{\mathbf{y}}}{(1-\varepsilon^2)n}$$

and note that

$$\begin{aligned}\sum_{i=1}^{d} D_i &\leqslant -\sum_{i=1}^{d} \frac{(\mathbf{t}_i^{\mathbf{x}})^2 + (\mathbf{t}_i^{\mathbf{y}})^2}{2n} + \sum_{i=1}^{d} \frac{d\varepsilon(M\mathbf{x})_i(M\mathbf{y})_i}{(1-\varepsilon^2)n} - \mathcal{N}\\ &\leqslant -\frac{1}{2n}\left(\sum_{i=1}^{d}(\mathbf{t}_i^{\mathbf{x}})^2 + \sum_{i=1}^{d}(\mathbf{t}_i^{\mathbf{y}})^2\right) + 2\delta^2|\varepsilon\langle M\mathbf{x}, M\mathbf{y}\rangle| + |\mathcal{N}|\end{aligned}$$

then holds, where in the last inequality we use the fact that $(1-\varepsilon^2) \geqslant 1/2$. Additionally,

$$\begin{aligned}|\mathcal{N}| &\leqslant \frac{\sqrt{d}}{(1-\varepsilon^2)n}\left|\sum_{i=1}^{d}\varepsilon\mathbf{s}_i^{\mathbf{y}}(M\mathbf{x})_i\right| + \frac{\sqrt{d}}{(1-\varepsilon^2)n}\left|\sum_{i=1}^{d}\varepsilon\mathbf{s}_i^{\mathbf{x}}(M\mathbf{y})_i\right| + \sum_{i=1}^{d}\frac{16|\varepsilon|}{(1-\varepsilon^2)n}\\ &\leqslant 2n^{-1}\sqrt{d}\,|\varepsilon\langle M\mathbf{x}, \mathbf{s}^{\mathbf{y}}\rangle| + 2n^{-1}\sqrt{d}\,|\varepsilon\langle M\mathbf{y}, \mathbf{s}^{\mathbf{x}}\rangle| + O\left(dn^{-1}\right)\\ &\leqslant 2n^{-1}\sqrt{d}\,|\varepsilon\langle M\mathbf{x}, \mathbf{s}^{\mathbf{y}}\rangle| + 2n^{-1}\sqrt{d}\,|\varepsilon\langle M\mathbf{y}, \mathbf{s}^{\mathbf{x}}\rangle| + o_d(1),\end{aligned}$$

where the second inequality holds since $|\varepsilon| \leqslant 1/2$ and thus $(1-\varepsilon^2) \geqslant 1/2$; the last inequality holds since $n = \omega(d)$ by the premise of the lemma; the constant 16 comes about from the fact that $s_i^{\mathbf{z}} \in [-4, 4]$ upheld by definition for every $i \in [d]$ and every $\mathbf{z} \in \{\mathbf{x}, \mathbf{y}\}$.

We conclude that

$$\sum_{i=1}^{d} D_i \leqslant -\frac{1}{2n}\left(\sum_{i=1}^{d}(\mathbf{t}_i^{\mathbf{x}})^2 + \sum_{i=1}^{d}(\mathbf{t}_i^{\mathbf{y}})^2\right) + 2\delta^2|\varepsilon\langle M\mathbf{x}, M\mathbf{y}\rangle|$$

$$+ 2n^{-1}\sqrt{d}\,|\varepsilon\langle M\mathbf{x}, \mathbf{s}^{\mathbf{y}}\rangle| + 2n^{-1}\sqrt{d}\,|\varepsilon\langle M\mathbf{y}, \mathbf{s}^{\mathbf{x}}\rangle| + o_d(1). \tag{22}$$

**Estimating $\sum \boldsymbol{E_i}$.** Start by writing

$$\sum_{i=1}^{d}\left(\frac{(k_i^{\mathbf{x}}+k_i^{\mathbf{y}})^3}{(\alpha n)^2}+\frac{(k_i^{\mathbf{x}}-k_i^{\mathbf{y}})^3}{((1-\alpha)n)^2}\right) \leqslant \frac{c_1\left|\sum_{i=1}^{d}(\mathbf{t}_i^{\mathbf{x}})^3\right| + c_2\left|\sum_{i=1}^{d}(\mathbf{t}_i^{\mathbf{y}})^3\right|}{n^2} + \frac{c_3\left|\sum_{i=1}^{d}(\mathbf{t}_i^{\mathbf{x}})^2\mathbf{t}_i^{\mathbf{y}}\right| + c_4\left|\sum_{i=1}^{d}\mathbf{t}_i^{\mathbf{x}}(\mathbf{t}_i^{\mathbf{y}})^2\right|}{n^2}, \tag{23}$$

where $c_1, c_2, c_3$, and $c_4$ are some constants, depending only on $\alpha$ (recall that $1/4 \leqslant \alpha \leqslant 3/4$ holds by Claim 10(a) and since $|\varepsilon| \leqslant 1/2$ by the premise of the lemma). As in (15), it follows by Lemma 9(ii) that

$$\left|\sum_{i=1}^{d}(\mathbf{t}_i^{\mathbf{x}})^3\right| \leqslant d^{3/2}\|M\mathbf{x}\|_\infty^2|\langle M\mathbf{x}, \mathbf{1}\rangle| + o(n^2) = O\left(d^{3/2}\log d\cdot|\langle M\mathbf{x}, \mathbf{1}\rangle|\right) + o(n^2)$$

and, similarly,

$$\left|\sum_{i=1}^{d}(\mathbf{t}_i^{\mathbf{y}})^3\right| \leqslant d^{3/2}\|M\mathbf{y}\|_\infty^2|\langle M\mathbf{y}, \mathbf{1}\rangle| + o(n^2) = O\left(d^{3/2}\log d\cdot|\langle M\mathbf{y}, \mathbf{1}\rangle|\right) + o(n^2).$$

An analogous argument shows that

$$\left|\sum_{i=1}^{d}(\mathbf{t}_i^{\mathbf{x}})^2\mathbf{t}_i^{\mathbf{y}}\right| \leqslant d^{3/2}\|M\mathbf{x}\|_\infty^2|\langle M\mathbf{y}, \mathbf{1}\rangle| + o(n^2) = O\left(d^{3/2}\log d\cdot|\langle M\mathbf{y}, \mathbf{1}\rangle|\right) + o(n^2)$$

and

$$\left|\sum_{i=1}^{d}\mathbf{t}_i^{\mathbf{x}}(\mathbf{t}_i^{\mathbf{y}})^2\right| \leqslant d^{3/2}\|M\mathbf{y}\|_\infty^2|\langle M\mathbf{x}, \mathbf{1}\rangle| + o(n^2) = O\left(d^{3/2}\log d\cdot|\langle M\mathbf{x}, \mathbf{1}\rangle|\right) + o(n^2).$$

It thus follows by (23) that

$$\begin{aligned}\sum_{i=1}^{d}E_i &= \sum_{i=1}^{d}\Theta\left(\frac{(k_i^{\mathbf{x}}+k_i^{\mathbf{y}})^3}{(\alpha n)^2}+\frac{(k_i^{\mathbf{x}}-k_i^{\mathbf{y}})^3}{((1-\alpha)n)^2}\right)\\ &= O\left(n^{-2}d^{3/2}\log d\right)\left(|\langle M\mathbf{x}, \mathbf{1}\rangle| + |\langle M\mathbf{y}, \mathbf{1}\rangle|\right) + o(1).\end{aligned} \tag{24}$$

**Estimating $\sum \boldsymbol{F_i}$.** It follows by (14) that

$$\left|\sum_{i=1}^{d}\mathbf{t}_i^{\mathbf{x}}\right| \leqslant \sqrt{d}|\langle M\mathbf{x}, \mathbf{1}\rangle| + o(n) \quad \text{and} \quad \left|\sum_{i=1}^{d}\mathbf{t}_i^{\mathbf{y}}\right| \leqslant \sqrt{d}|\langle M\mathbf{y}, \mathbf{1}\rangle| + o(n).$$

both hold. Hence,

$$\sum_{i=1}^{d} F_i = \sum_{i=1}^{d} o\left(\frac{k_i^{\mathbf{x}} + k_i^{\mathbf{y}}}{\alpha n} + \frac{k_i^{\mathbf{x}} - k_i^{\mathbf{y}}}{(1-\alpha)n}\right) = o\left(\frac{\left|\sum_{i=1}^{d} \mathbf{t}_i^{\mathbf{x}}\right|}{n}\right) + o\left(\frac{\left|\sum_{i=1}^{d} \mathbf{t}_i^{\mathbf{y}}\right|}{n}\right)$$
$$= o\left(n^{-1}\sqrt{d}\right)\left(|\langle M\mathbf{x}, \mathbf{1}\rangle| + |\langle M\mathbf{y}, \mathbf{1}\rangle|\right) + o(1). \tag{25}$$

We conclude the proof by noticing that combining (22), (24), and (25) implies (20). □

Following Lemma 23, we turn to estimate the terms $\exp(-\delta_{\mathbf{x}} - \delta_{\mathbf{y}})$ and $\beta(\mathbf{x}, \mathbf{y})$; in fact, estimations of these in expectation suffices. In the sequel, we prove the following two lemmas.

**Lemma 25.** *If $n = \omega(d \log d)$, then*

$$\mathbb{E}_{\mathbf{x}\sim\mathcal{S}}\left[\exp\left(\delta_{\mathbf{x}}\right)\right] = 1 + o_d(1).$$

**Lemma 26.** *If $n = \omega(d \log d)$, then*

$$\mathbb{E}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}[\beta(\mathbf{x}, \mathbf{y})] = 1 + o_d(1).$$

Postponing the proofs of Lemmas 25 and 26 until later, we first deduce (12) from these and thus conclude the proof of Claim 17; this deduction is captured in the following lemma.

**Lemma 27.** *If $n = \omega(d \log d)$, then*

$$\mathbb{E}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}[P_{\mathbf{x},\mathbf{y}}] \leqslant (1 + o_d(1))\mathbb{E}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}[P_{\mathbf{x}}P_{\mathbf{y}}].$$

*Proof.* First, note that

$$\begin{aligned}
\mathbb{E}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}[P_{\mathbf{x}}P_{\mathbf{y}}] &= \mathbb{E}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}\left[(1 + o_d(1))p^2 \exp\left(\delta_{\mathbf{x}} + \delta_{\mathbf{y}}\right)\right] \\
&= (1 + o_d(1))p^2 \cdot \mathbb{E}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}\left[\exp\left(\delta_{\mathbf{x}}\right) \cdot \exp\left(\delta_{\mathbf{y}}\right)\right] \\
&= (1 + o_d(1))p^2 \cdot \mathbb{E}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}\left[\exp\left(\delta_{\mathbf{x}}\right)\right] \cdot \mathbb{E}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}\left[\exp\left(\delta_{\mathbf{y}}\right)\right] \\
&\geqslant (1 - o_d(1))p^2,
\end{aligned}$$

where the first equality holds by Lemma 21, the last equality holds since $\mathbf{x}$ and $\mathbf{y}$ are sampled independently and thus $\exp(\delta_{\mathbf{x}})$ and $\exp(\delta_{\mathbf{y}})$ are independent, and the inequality holds by Lemma 25.

Next, let $\mathcal{E}$ denote the event that $|\varepsilon| > 1/2$. Since $p$ is fixed, it follows that

$$\begin{aligned}
\mathbb{E}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}[P_{\mathbf{x},\mathbf{y}}] &\leqslant \mathbb{E}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}[P_{\mathbf{x},\mathbf{y}}|\bar{\mathcal{E}}] + \mathbb{E}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}[P_{\mathbf{x},\mathbf{y}}|\mathcal{E}] \cdot \mathbb{P}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}[\mathcal{E}] \\
&\leqslant \mathbb{E}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}[P_{\mathbf{x},\mathbf{y}}|\bar{\mathcal{E}}] + \mathbb{E}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}[(1 + o_d(1))p \cdot \exp(\delta_{\mathbf{x}})|\ \mathcal{E}] \cdot \mathbb{P}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}[\mathcal{E}] \\
&= \mathbb{E}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}[P_{\mathbf{x},\mathbf{y}}|\bar{\mathcal{E}}] + (1 + o_d(1))p^2 \cdot p^{-1} \cdot \mathbb{E}_{\mathbf{x}\sim\mathcal{S}}[\exp(\delta_{\mathbf{x}})] \cdot \mathbb{P}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}[\mathcal{E}]
\end{aligned}$$

$$\leqslant \mathbb{E}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}[P_{\mathbf{x},\mathbf{y}}|\bar{\mathcal{E}}] + (1+o_d(1))p^2 \cdot p^{-1} \cdot \mathbb{P}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}[\mathcal{E}] \tag{26}$$

where the second inequality holds since $P_{\mathbf{x},\mathbf{y}} \leqslant \min\{P_{\mathbf{x}}, P_{\mathbf{y}}\} \leqslant (1+o_d(1))p \cdot \exp(\delta_{\mathbf{x}})$ by Lemma 21, the equality holds since, once $\mathbf{y}$ becomes irrelevant, so does the event $\mathcal{E}$, and the last inequality holds by Lemma 25.

Given any vector $\mathbf{y} \in \{-1,1\}^n$, note that $n^{-1/2}\mathbf{y} \in \mathbb{S}^{n-1}$. It thus follows by Lemma 9(iii) that

$$\begin{aligned}\mathbb{P}_{\mathbf{x}\sim\mathcal{S}}\left[|n^{-1}\langle \mathbf{x},\mathbf{y}\rangle| > 1/2\right] &= \mathbb{P}_{\mathbf{x}\sim\mathcal{S}}\left[\left|\left\langle \mathbf{x}, n^{-1/2}\mathbf{y}\right\rangle\right| > \sqrt{n}/2\right] \\ &\leqslant d^{C_9} \cdot \exp(-n/36) \leqslant \exp(-n/40),\end{aligned}$$

holds for every $\mathbf{y} \in \{-1,1\}^n$. It then follows by the law of total probability that

$$\mathbb{P}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}[\mathcal{E}] = \mathbb{P}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}[|\varepsilon| > 1/2] \leqslant \exp(-\Theta(n)). \tag{27}$$

On the other hand, since $p = \left(\frac{8}{\pi n}\right)^{d/2} \cdot \exp(-r^2\delta^2/2)$ and $n = \omega(d\log d)$, it follows that

$$p^{-1} = \left(\frac{\pi n}{8}\right)^{d/2} \cdot \exp(r^2\delta^2/2) \leqslant \exp\left(O(d\log n) + O\left(d^2/n\right)\right) = \exp(o(n)). \tag{28}$$

It then follows by (26), (27) and (28) that

$$\mathbb{E}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}[P_{\mathbf{x},\mathbf{y}}] \leqslant \mathbb{E}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}[P_{\mathbf{x},\mathbf{y}}|\bar{\mathcal{E}}] + p^2 \cdot o_n(1). \tag{29}$$

Conditioning on $\bar{\mathcal{E}}$ and recalling that $n = \omega(d\log d)$ holds by the premise of the lemma, we obtain

$$P_{\mathbf{x},\mathbf{y}} \leqslant (1+o_d(1))P_{\mathbf{x}}P_{\mathbf{y}} \cdot \beta(\mathbf{x},\mathbf{y}) \cdot \exp(-\delta_{\mathbf{x}} - \delta_{\mathbf{y}}) = (1+o_d(1)) \cdot p^2 \cdot \beta(\mathbf{x},\mathbf{y}), \tag{30}$$

where the inequality holds by Lemma 23 and the equality holds by Lemma 21.

Combining (29) and (30) we conclude that

$$\begin{aligned}\mathbb{E}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}[P_{\mathbf{x},\mathbf{y}}] &\leqslant \mathbb{E}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}[(1+o_d(1)) \cdot p^2 \cdot \beta(\mathbf{x},\mathbf{y})|\bar{\mathcal{E}}] + p^2 \cdot o_n(1) \\ &\leqslant (1+o_d(1)) \cdot p^2 \cdot \mathbb{E}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}[\beta(\mathbf{x},\mathbf{y})|\bar{\mathcal{E}}] + p^2 \cdot o_n(1) \\ &\leqslant (1+o_d(1)) \cdot p^2 \cdot \mathbb{E}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}[\beta(\mathbf{x},\mathbf{y})] + p^2 \cdot o_n(1) \\ &= (1+o_d(1)) \cdot p^2 + p^2 \cdot o_n(1) \\ &= (1+o_d(1)) \cdot p^2,\end{aligned}$$

where the last inequality holds since $\mathbb{P}[\bar{\mathcal{E}}] = 1 - o(1)$ and $\beta$ is non-negative, and the first equality holds by Lemma 26. □

It remains to prove Lemmas 25 and 26. The following result facilitates our proofs of these lemmas; proof of the latter is essentially that seen for [16, Lemma 2.4] and is thus omitted.

**Lemma 28.** [16, Lemma 2.4] *Let $X$ be a non-negative random variable which satisfies*

$$\mathbb{P}[X > t] \leqslant d^{\xi_1} \cdot e^{-t^2/\xi_2} \text{ for any } t > 0,$$

*for some positive constants $\xi_1$ and $\xi_2$. Then, for any $\lambda = \kappa\sqrt{\log d}$, where $\kappa \geqslant 2\sqrt{\xi_1\xi_2}$, we have*

$$\mathbb{E}\left[\exp\left(X^2/\lambda^2\right)\right] \leqslant 1 + 4\xi_1\xi_2/\kappa^2 + o_d(1).$$

*Proof of Lemma 25.* Starting with the upper bound, given any $\mathbf{x} \in \mathcal{S}$, let

$$Z := Z(\mathbf{x}) = \left(cn^{-2}d^{3/2}\log d + \zeta n^{-1}\sqrt{d}\right)|\langle M\mathbf{x}, \mathbf{1}\rangle|,$$

where $c$ is a constant and $\zeta = o(1)$ are chosen so as to ensure that $|\delta_{\mathbf{x}}| \leqslant Z$ holds. If $\left|\left\langle M\mathbf{x}, d^{-1/2}\mathbf{1}\right\rangle\right| < 1$, then the assumption that $n = \omega(d\log d)$ implies that $Z = o_d(1)$ and thus

$$\mathbb{E}_{\mathbf{x}\sim\mathcal{S}}\left[\exp\left(\delta_{\mathbf{x}}\right)\right] \leqslant \mathbb{E}_{\mathbf{x}\sim\mathcal{S}}\left[\exp\left(Z\right)\right] \leqslant 1 + o_d(1).$$

Assume then that $\left|\left\langle M\mathbf{x}, d^{-1/2}\mathbf{1}\right\rangle\right| \geqslant 1$. In this case

$$Z = \left(cn^{-2}d^2\log d + \zeta n^{-1}\sqrt{d}\right)\left|\left\langle M\mathbf{x}, d^{-1/2}\mathbf{1}\right\rangle\right| \leqslant \left(cn^{-2}d^2\log d + \zeta n^{-1}\sqrt{d}\right)\left|\left\langle M\mathbf{x}, d^{-1/2}\mathbf{1}\right\rangle\right|^2.$$

Writing $\gamma := \left\langle M\mathbf{x}, d^{-1/2}\mathbf{1}\right\rangle$ and $\lambda := \sqrt{n/d}$, it then follows that

$$Z \leqslant \left(cn^{-2}d^2\log d + \zeta n^{-1}d\right)|\gamma|^2 \leqslant n^{-1}d|\gamma|^2 = |\gamma|^2/\lambda^2,$$

where the second inequality holds since $\zeta = o(1)$ and since $n = \omega(d\log d)$ implies that $n^{-2}d^2\log d = o\left(n^{-1}d\right)$.

Then,

$$\mathbb{P}_{\mathbf{x}\sim\mathcal{S}}[|\gamma| > t] = \mathbb{P}_{\mathbf{x}\sim\mathcal{S}}\left[|\langle M\mathbf{x}, d^{-1/2}\mathbf{1}\rangle| > t\right] \leqslant d^{C_9}\exp\left(-t^2/8\right)$$

holds for any $t > 0$, by Lemma 9(iii). We conclude that the non-negative random variable $X := |\gamma|$ satisfies the conditions of Lemma 28, implying that

$$\mathbb{E}_{\mathbf{x}\sim\mathcal{S}}\left[\exp\left(\delta_{\mathbf{x}}\right)\right] \leqslant \mathbb{E}_{\mathbf{x}\sim\mathcal{S}}\left[\exp\left(|\delta_{\mathbf{x}}|\right)\right] \leqslant \mathbb{E}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}[\exp(Z)] \leqslant \mathbb{E}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}\left[\exp\left(X^2/\lambda^2\right)\right] = 1 + o_d(1), \tag{31}$$

where the last equality holds since $n = \omega(d\log d)$ by the premise of the lemma and thus $\lambda = \omega(\sqrt{\log d})$.

Next, we prove that $\mathbb{E}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}\left[\exp\left(\delta_{\mathbf{x}}\right)\right] \geqslant 1 - o_d(1)$. Let $Y = \exp\left(|\delta_{\mathbf{x}}|\right)$; note that $Y$ is positive. Let $g : \mathbb{R}^+ \to \mathbb{R}^+$ be defined by $g(x) = 1/x$; note that $g$ is convex. It thus follows by Jensen's inequality[11] that

$$\begin{aligned}\mathbb{E}_{\mathbf{x}\sim\mathcal{S}}\left[\exp\left(\delta_{\mathbf{x}}\right)\right] &\geqslant \mathbb{E}_{\mathbf{x}\sim\mathcal{S}}\left[\exp\left(-|\delta_{\mathbf{x}}|\right)\right] = \mathbb{E}_{\mathbf{x}\sim\mathcal{S}}[g(Y)] \geqslant g\left(\mathbb{E}_{\mathbf{x}\sim\mathcal{S}}[Y]\right)\\ &= \frac{1}{\mathbb{E}_{\mathbf{x}\sim\mathcal{S}}\left[\exp\left(|\delta_{\mathbf{x}}|\right)\right]} \geqslant \frac{1}{1 + o_d(1)} = 1 - o_d(1),\end{aligned}$$

where the last inequality holds by (31). □

[11] Jensen's inequality, see, e.g. [30], asserts that $g(\mathbb{E}(X)) \leqslant \mathbb{E}(g(X))$ holds whenever $X$ is a random variable and $g$ is a convex function.

*Proof of Lemma 26.* Let

$$Z_1 := Z_1(\mathbf{x},\mathbf{y}) = d\varepsilon^2 + 2\delta^2|\varepsilon\langle M\mathbf{x}, M\mathbf{y}\rangle| + 2n^{-1}\sqrt{d}\,|\varepsilon\langle M\mathbf{x}, \mathbf{s}^{\mathbf{y}}\rangle| + 2n^{-1}\sqrt{d}\,|\varepsilon\langle M\mathbf{y}, \mathbf{s}^{\mathbf{x}}\rangle|$$

and let

$$Z_2 := Z_2(\mathbf{x},\mathbf{y}) = \left(cn^{-2}d^{3/2}\log d + \zeta n^{-1}\sqrt{d}\right)\left(|\langle M\mathbf{x}, \mathbf{1}\rangle| + |\langle M\mathbf{y}, \mathbf{1}\rangle|\right),$$

where $c$ is a constant and $\zeta = o(1)$.

Set

$$\begin{aligned}
\bar{\varepsilon} &:= \bar{\varepsilon}(\mathbf{x},\mathbf{y}) = \left\langle \mathbf{x}, n^{-1/2}\mathbf{y}\right\rangle,\\
\mu &:= \mu(\mathbf{x},\mathbf{y}) = \left\langle M\mathbf{x}, \|M\mathbf{y}\|_2^{-1}M\mathbf{y}\right\rangle,\\
\eta_1 &:= \eta_1(\mathbf{x},\mathbf{y}) = \left\langle M\mathbf{x}, \|\mathbf{s}^{\mathbf{y}}\|_2^{-1}\mathbf{s}^{\mathbf{y}}\right\rangle,\\
\eta_2 &:= \eta_2(\mathbf{x},\mathbf{y}) = \left\langle M\mathbf{y}, \|\mathbf{s}^{\mathbf{x}}\|_2^{-1}\mathbf{s}^{\mathbf{x}}\right\rangle.
\end{aligned}$$

Recalling that $\delta = \sqrt{d/n}$, we obtain

$$\begin{aligned}
Z_1 &= \frac{d}{n}\cdot|\bar{\varepsilon}|^2 + 2\frac{d\|M\mathbf{y}\|_2}{n^{3/2}}|\bar{\varepsilon}||\mu| + 2\frac{\sqrt{d}\|\mathbf{s}^{\mathbf{y}}\|_2}{n^{3/2}}|\bar{\varepsilon}||\eta_1| + 2\frac{\sqrt{d}\|\mathbf{s}^{\mathbf{x}}\|_2}{n^{3/2}}|\bar{\varepsilon}||\eta_2|\\
&\leqslant \frac{d}{n}\cdot|\bar{\varepsilon}|^2 + O\left(\frac{d^{3/2}}{n^{3/2}}\right)|\bar{\varepsilon}||\mu| + 8\frac{d}{n^{3/2}}|\bar{\varepsilon}||\eta_1| + 8\frac{d}{n^{3/2}}|\bar{\varepsilon}||\eta_2|\\
&\leqslant \left(\frac{d}{n} + O\left(\frac{d^{3/2}}{n^{3/2}}\right) + \frac{16d}{n^{3/2}}\right)\left(|\bar{\varepsilon}| + |\mu| + |\eta_1| + |\eta_2|\right)^2\\
&\leqslant 2n^{-1}d(|\bar{\varepsilon}| + |\mu| + |\eta_1| + |\eta_2|)^2, \qquad (32)
\end{aligned}$$

where the first inequality holds by Lemma 9(i) and by Lemma 13(1), and the last inequality holds since $n = \omega(d)$.

Let $\gamma = \left|\left\langle M\mathbf{x}, d^{-1/2}\mathbf{1}\right\rangle\right| + \left|\left\langle M\mathbf{y}, d^{-1/2}\mathbf{1}\right\rangle\right|$. If $\gamma < 1$, then $Z_2 = o_d(1)$; otherwise

$$\begin{aligned}
Z_2 &= \left(cn^{-2}d^2\log d + \zeta n^{-1}d\right)\left(\left|\left\langle M\mathbf{x}, d^{-1/2}\mathbf{1}\right\rangle\right| + \left|\left\langle M\mathbf{y}, d^{-1/2}\mathbf{1}\right\rangle\right|\right)\\
&= O\left(n^{-2}d^2\log d + \zeta n^{-1}d\right)\gamma^2 \leqslant n^{-1}d\gamma^2, \qquad (33)
\end{aligned}$$

where the last inequality holds since $\zeta = o(1)$ and since $n = \omega(d\log d)$ implies that $n^{-2}d^2\log d = o\left(n^{-1}d\right)$.

Let $\lambda := \sqrt{n/(2d)}$. Combining (32) and (33) we obtain

$$\begin{aligned}
Z_1 + Z_2 &\leqslant 2n^{-1}d(|\bar{\varepsilon}| + |\mu| + |\eta_1| + |\eta_2|)^2 + n^{-1}d\gamma^2 \leqslant 2n^{-1}d(|\bar{\varepsilon}| + |\mu| + |\eta_1| + |\eta_2| + \gamma)^2\\
&\leqslant (|\bar{\varepsilon}| + |\mu| + |\eta_1| + |\eta_2| + \gamma)^2/\lambda^2. \qquad (34)
\end{aligned}$$

Given any vector $\mathbf{y} \in \{-1,1\}^n$, note that $n^{-1/2}\mathbf{y} \in \mathbb{S}^{n-1}$. It thus follows by Lemma 9(iii) that

$$\mathbb{P}_{\mathbf{x}\sim\mathcal{S}}\left[\left|\left\langle \mathbf{x}, n^{-1/2}\mathbf{y}\right\rangle\right| > t\right] \leqslant d^{C_9}\exp\left(-t^2/9\right)$$

holds for any $t > 0$. It then follows by the law of total probability that

$$\mathbb{P}[|\bar{\varepsilon}| > t] = \mathbb{P}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}\left[|\langle \mathbf{x}, n^{-1/2}\mathbf{y}\rangle| > t\right] \leqslant d^{C_9}\exp\left(-t^2/9\right) \text{ for any } t > 0. \tag{35}$$

Similarly, since $\|M\mathbf{y}\|_2^{-1}M\mathbf{y} \in \mathbb{S}^{d-1}$ and $\|\mathbf{s}^{\mathbf{y}}\|_2^{-1}\mathbf{s}^{\mathbf{y}} \in \mathbb{S}^{d-1}$ hold for every $\mathbf{y} \in \operatorname{Supp}\mathcal{S}$, $\|\mathbf{s}^{\mathbf{x}}\|_2^{-1}\mathbf{s}^{\mathbf{x}} \in \mathbb{S}^{d-1}$ holds for every $\mathbf{x} \in \operatorname{Supp}\mathcal{S}$, and $d^{-1/2}\mathbf{1} \in \mathbb{S}^{d-1}$, it follows that

$$\begin{aligned}
\mathbb{P}[|\mu| > t] &\leqslant d^{C_9}\exp\left(-t^2/9\right) \text{ for any } t > 0\\
\mathbb{P}[|\eta_1| > t] &\leqslant d^{C_9}\exp\left(-t^2/9\right) \text{ for any } t > 0\\
\mathbb{P}[|\eta_2| > t] &\leqslant d^{C_9}\exp\left(-t^2/9\right) \text{ for any } t > 0\\
\mathbb{P}[\gamma > t] &\leqslant d^{C_9}\exp\left(-t^2/9\right) \text{ for any } t > 0.
\end{aligned}$$

We conclude that the non-negative random variable $X := |\bar{\varepsilon}| + |\mu| + |\eta_1| + |\eta_2| + \gamma$ satisfies the conditions of Lemma 28, implying that

$$\mathbb{E}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}[\beta(\mathbf{x},\mathbf{y})] = \mathbb{E}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}[\exp(Z_1 + Z_2)] \leqslant \mathbb{E}_{\mathbf{x},\mathbf{y}\sim\mathcal{S}}\left[\exp\left(X^2/\lambda^2\right)\right] = 1 + o_d(1),$$

where the last equality holds since $n = \omega(d\log d)$ by the premise of the lemma and thus $\lambda = \omega(\sqrt{\log d})$. □

## 4 Concluding remarks

We have proved that $\mathrm{DISC}(M + R/\sqrt{d}) = O(d^{-1/2})$ holds asymptotically almost surely, whenever $M \in \mathbb{R}^{d\times n}$ is Komlós, $R \in \mathbb{R}^{d\times n}$ is Rademacher, $d = \omega(1)$, and $n = \omega(d\log d)$.

As stated by Bansal, Jiang, Meka, Singla, and Sinha [16, Section 3], considering other distributions for the entries of the random perturbation and specifically discrete ones is of high interest as well. In view of the aforementioned result in [11] pertaining to the discrepancy of Bernoulli matrices, as well as the proclaimed $n = \omega(d^2)$ bound attained in [16] in the smoothed setting with Bernoulli noise, the following question seems to be a natural next step.

**Question 29.** Let $d = \omega(1)$ and $n = \omega(d\log d)$ be integers, and set $p := p(n,d) > 0$. Is it true that $\mathrm{DISC}(M + R) = O(1)$ holds a.a.s. whenever $M \in \mathbb{R}^{d\times n}$ is a Komlós matrix and $R \in \mathbb{R}^{d\times n}$ is a random matrix with each of its entries being an independent copy of $\Theta\left((pd)^{-1/2}\right)\mathrm{Ber}(p)$?

It is conceivable that for certain ranges of $p$, the $O(1)$ bound on the discrepancy, appearing in Question 29, can be replaced with $1/\mathrm{poly}(d)$.

### Acknowledgements

We would like to express our gratitude to an anonymous referee providing us with helpful comments.

## References


[1] E. Abbe, S. Li, and A. Sly. Proof of the contiguity conjecture and lognormal limit for the symmetric perceptron, in *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science*, *IEEE Computer Society*, Los Alamitos, CA, 2022, pp. 327–338.

[2] E. Aigner-Horev, O. Danon, D. Hefetz, and S. Letzter. Large rainbow cliques in randomly perturbed dense graphs. *SIAM Journal on Discrete Mathematics*, 36: 2975–2994, 2022.

[3] E. Aigner-Horev, O. Danon, D. Hefetz, and S. Letzter. Small rainbow cliques in randomly perturbed dense graphs. *European Journal of Combinatorics*, 101: 103452, 2022.

[4] E. Aigner-Horev and D. Hefetz. Rainbow Hamilton cycles in randomly coloured randomly perturbed dense graphs. *SIAM Journal on Discrete Mathematics*, 35: 1569–1577, 2021.

[5] E. Aigner-Horev, D. Hefetz, and K. Krivelevich. Cycle lengths in randomly perturbed graphs. *Random Structures & Algorithms*, 63: 867–884, 2023.

[6] E. Aigner-Horev, D. Hefetz, and K. Krivelevich. Minors, connectivity, and diameter in randomly perturbed sparse graphs. arXiv:2212.07192, 2022.

[7] E. Aigner-Horev, D. Hefetz, and A. Lahiri. Rainbow trees in uniformly edge-coloured graphs. *Random Structures & Algorithms*, 62: 287–303, 2023.

[8] E. Aigner-Horev, D. Hefetz, and M. Schacht. Ramsey properties of randomly perturbed hypergraphs. arXiv:2311.01750, 2023.

[9] E. Aigner-Horev and Y. Person. Monochromatic Schur triples in randomly perturbed dense sets of integers. *SIAM Journal on Discrete Mathematics*, 33: 2175–2180, 2019.

[10] E. Aigner-Horev and Y. Person. On sparse random combinatorial matrices. *Discrete Mathematics*, 345: 113017, 2022.

[11] D. J. Altschuler and J. Niles-Weed. The discrepancy of random rectangular matrices, *Random Structures & Algorithms*, 60: 551–593, 2022.

[12] B. Aubin, W. Perkins, and L. Zdeborová. Storage capacity in symmetric binary perceptrons, *Journal of Physics. A. Mathematical and Theoretical*, 52: 294003, 2019.

[13] J. Balogh, A. Treglown, and A. Z. Wagner. Tilings in randomly perturbed dense graphs. *Combinatorics, Probability and Computing*, 28: 159–176, 2019.

[14] W. Banaszczyk. Balancing vectors and Gaussian measures of $n$-dimensional convex bodies, *Random Structures & Algorithms*, 12: 351–360, 1998.

[15] N. Bansal. Constructive algorithms for discrepancy minimization, in *2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, *IEEE Computer Society*, Los Alamitos, CA, 2010, pp. 3–10.

[16] N. Bansal, H. Jiang, R. Meka, S. Singla, and M. Sinha. Smoothed analysis of the Komlós conjecture. In *49th EATCS International Conference on Automata, Languages, and Programming*, volume 229 of *LIPIcs. Leibniz Int. Proc. Inform.*, Paper No. 14, 12, Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2022.

[17] J. Beck and T. Fiala. "Integer-making" theorems. *Discrete Applied Mathematics. The Journal of Combinatorial Algorithms, Informatics and Computational Sciences*, 3: 1–8, 1981.

[18] J. Beck and V. T. Sś. Discrepancy theory, in *Handbook of combinatorics, Vol. 1, 2*, Elsevier Sci. B. V., Amsterdam, 1995, pp. 1405–1446.

[19] W. Bedenknecht, J. Han, Y. Kohayakawa, and G. O. Mota. Powers of tight Hamilton cycles in randomly perturbed hypergraphs. *Random Structures & Algorithms*, 55: 795–807, 2019.

[20] T. Bohman, A. Frieze, M. Krivelevich, and R. Martin. Adding random edges to dense graphs. *Random Structures & Algorithms*, 24: 105–117, 2004.

[21] T. Bohman, A. Frieze, and R. Martin. How many random edges make a dense graph Hamiltonian? *Random Structures & Algorithms*, 22: 33–42, 2003.

[22] B. Bukh. An improvement of the Beck–Fiala theorem. *Combinatorics, Probability and Computing*, 25: 380–398, 2016.

[23] J. Böttcher, R. Montgomery, O. Parczyk, and Y. Person. Embedding spanning bounded degree graphs in randomly perturbed graphs. *Mathematika*, 66: 422–447, 2020.

[24] J. Böttcher, J. Han, Y. Kohayakawa, R. Montgomery, O. Parczyk, and Y. Person. Universality for bounded degree spanning trees in randomly perturbed graphs. *Random Structures & Algorithms*, 55: 854–864, 2019.

[25] K. Chandrasekaran and S. S. Vempala. Integer feasibility of random polytopes, in *ITCS'14—Proceedings of the 2014 Conference on Innovations in Theoretical Computer Science*, ACM, New York, 2014, pp. 449–458.

[26] M. Charikar, A. Newman, and A. Nikolov. Tight hardness results for minimizing discrepancy, in *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms*, SIAM, Philadelphia, PA, 2011, pp. 1607–1614.

[27] B. Chazelle. The discrepancy method: Randomness and complexity. Cambridge University Press, 2000, pp. xviii+463.

[28] K. P. Costello. Balancing Gaussian vectors, *Israel Journal of Mathematics*, 172: 145–156, 2009.

[29] A. Dudek, C. Reiher, A. Ruciński, and M. Schacht. Powers of Hamiltonian cycles in randomly augmented graphs. *Random Structures & Algorithms*, 56: 122–141, 2020.

[30] R. Durrett. Probability—theory and examples, Fifth edition, *Cambridge Series in Statistical and Probabilistic Mathematics*, volume 49, *Cambridge University Press*, Cambridge, 2019, xii+419.

[31] E. Ezra and S. Lovett. On the Beck-Fiala conjecture for random set systems, *Random Structures & Algorithms*, 54: 665–675, 2019.

[32] J. Han and Y. Zhao. Hamiltonicity in randomly perturbed hypergraphs. *Journal of Combinatorial Theory Series B*, 144: 14–31, 2020.

[33] R. Harishchandra, A. Levy, and T. Rothvoss. Deterministic Discrepancy Minimization via the Multiplicative Weight Update Method, in *Integer Programming and Combinatorial Optimization - 19th International Conference, IPCO 2017, Waterloo, ON, Canada, June 26-28, 2017, Proceedings*, volume 10328 of *Lecture Notes in Computer Science*, Springer, 2017, pp. 380–391.

[34] C. Harshaw, F. Sävje, D. Spielman, and P. Zhang. Balancing covariates in randomized experiments with the Gram-Schmidt Walk design, *Journal of the American Statistical Association*, 2024.

[35] R. Hoberg and T. Rothvoss. A Fourier-analytic approach for the discrepancy of random set systems, in *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms*, SIAM, Philadelphia, PA, 2019, pp. 2547–2556.

[36] M. Krivelevich, M. Kwan, and B. Sudakov. Bounded-Degree Spanning Trees in Randomly Perturbed Graphs. *SIAM Journal on Discrete Mathematics*, 31: 155–171, 2017.

[37] M. Krivelevich, M. Kwan, and B. Sudakov. Cycles and matchings in randomly perturbed digraphs and hypergraphs. *Combinatorics, Probability and Computing*, 25: 909–927, 2016.

[38] M. Krivelevich, B. Sudakov, and P. Tetali. On smoothed analysis in dense graphs and formulas. *Random Structures & Algorithms*, 29: 180–193, 2006.

[39] S. Lovett and R. Meka. Constructive discrepancy minimization by walking on the edges, *SIAM Journal on Computing*, 44: 1573–1582, 2015.

[40] A. McDowell and R. Mycroft. Hamilton $\ell$-cycles in randomly perturbed hypergraphs. *Electronic Journal of Combinatorics*, 25: Paper 4.36, 30, 2018.

[41] R. Meka, P. Rigollet, and P. Turner. Balancing Gaussian vectors in high dimension, in *Conference on Learning Theory, COLT 2020, 9-12 July 2020, Virtual Event [Graz, Austria]*, *Proceedings of Machine Learning Research, volume 125*, PMLR, 2020, pp. 3455–3486.

[42] W. Perkins and C. Xu. Frozen 1-RSB structure of the symmetric Ising perceptron, in *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, ACM, New York, 2021, pp. 1579–1588.

[43] A. Potukuchi. Discrepancy in random hypergraph models. arXiv:1811.01491, 2018.

[44] T. Rothvoss. Constructive discrepancy minimization for convex sets, *SIAM Journal on Computing*, 46: 224–234, 2017.

[45] J. Spencer. Six standard deviations suffice, Transactions of the American Mathematical Society, 289: 679–706, 1985.

[46] J. Spencer and L. Florescu. Asymptopia, *Student Mathematical Library, vol. 71, American Mathematical Society*, 2014, xiv+183.

[47] D. Spielman and S. Teng. Smoothed analysis: an attempt to explain the behavior of algorithms in practice, *Communications of the ACM*, 52: 76–84, 2009.

[48] T. Tao and V. Vu. Random Matrices: The circular law, *Communications in Contemporary Mathematics*, 10: 261–307, 2008.

[49] T. Tao and V. Vu. Smoothed analysis of the condition number and the least singular value, *Mathematics of Computation*, 79: 2333–2352, 2010.

[50] T. Tao and V. Vu. The condition number of a randomly perturbed matrix, in *STOC'07—Proceedings of the 39th Annual ACM Symposium on Theory of Computing*, ACM, New York, 2007, pp. 248–255.

[51] R. Vershynin. High-dimensional probability, *Cambridge Series in Statistical and Probabilistic Mathematics*, vol. 47, Cambridge University Press, Cambridge, 2018, xiv+284.

# A Proofs of Lemmas 15 and 16

Prior to proving Lemmas 15 and 16, we collect several auxiliary results.

**Observation 30.** *Let $n \in \mathbb{N}$ be even and let $\mathbf{x}, \mathbf{y} \in \{-1,1\}^n$ satisfying $\#_1(\mathbf{x}) \equiv \#_1(\mathbf{y}) \pmod 2$ be given. Then, $|\mathrm{Diff}(\mathbf{x},\mathbf{y})|$ is even.*

*Proof.* Let $A := A(\mathbf{x},\mathbf{y}) = |\{i \in [n] : \mathbf{x}_i = \mathbf{y}_i = 1\}|$, let $B := B(\mathbf{x},\mathbf{y}) = |\{i \in [n] : \mathbf{x}_i = \mathbf{y}_i = -1\}|$, let $C := C(\mathbf{x},\mathbf{y}) = |\{i \in [n] : \mathbf{x}_i = 1, \mathbf{y}_i = -1\}|$, and let $D := D(\mathbf{x},\mathbf{y}) = |\{i \in [n] : \mathbf{x}_i = -1, \mathbf{y}_i = 1\}|$. Suppose for a contradiction that $|\mathrm{Diff}(\mathbf{x},\mathbf{y})|$ is odd. Since $|\mathrm{Diff}(\mathbf{x},\mathbf{y})| = C + D$, we may assume without loss of generality that $C$ is even and $D$ is odd. Since, moreover, $n = A + B + C + D$ is even, we may further assume without loss of generality that $A$ is even and $B$ is odd. It then follows that $\#_1(\mathbf{x}) = A + C$ is even, whereas $\#_1(\mathbf{y}) = A + D$ is odd; this contradicts the premise of the observation and concludes its proof. □

**Lemma 31.** *Let $n \in \mathbb{N}$, let $k \in \mathbb{Z}$, and let $\mathbf{u}, \mathbf{v} \in \{-1,1\}^n$ be vectors satisfying $\sum_{i=1}^n \mathbf{u}_i = 2k = \sum_{i=1}^n \mathbf{v}_i$. Then, $|\mathrm{Diff}(\mathbf{v},\mathbf{u})|$ is even.*

*Proof.* Set

$$O = \{i \in \mathrm{Diff}(\mathbf{v},\mathbf{u}) : \mathbf{u}_i = 1\} \text{ and } M = \{i \in \mathrm{Diff}(\mathbf{v},\mathbf{u}) : \mathbf{u}_i = -1\}.$$

Then

$$\begin{aligned} 2k = \sum_{i=1}^n \mathbf{v}_i &= \sum_{i \notin \mathrm{Diff}(\mathbf{v},\mathbf{u})} \mathbf{u}_i + \sum_{i \in O} (\mathbf{u}_i - 2) + \sum_{i \in M} (\mathbf{u}_i + 2) \\ &= \sum_{i=1}^n \mathbf{u}_i - 2|O| + 2|M| = 2k - 2|O| + 2|M|. \end{aligned}$$

It follows that $|O| = |M|$, and thus $|\mathrm{Diff}(\mathbf{v},\mathbf{u})| = |O| + |M|$ is even. □

**Lemma 32.** *Let $\mathbf{u} \in \{-1,1\}^n$ and let $\mathbf{v} \in \mathcal{E}_n$. If $|\mathrm{Diff}(\mathbf{v},\mathbf{u})|$ is even, then $\mathbf{u} \in \mathcal{E}_n$.*

*Proof.* The proof is via induction on $|\mathrm{Diff}(\mathbf{v},\mathbf{u})|$. If $|\mathrm{Diff}(\mathbf{v},\mathbf{u})| = 0$, then $\mathbf{u} = \mathbf{v} \in \mathcal{E}_n$. Suppose then that $|\mathrm{Diff}(\mathbf{v},\mathbf{u})| = 2$ and let $i, j \in [n]$ be the (sole) two distinct indices over which $\mathbf{u}$ and $\mathbf{v}$ differ. The equality $\#_1(\mathbf{u}) = \#_1(\mathbf{v}) - (\mathbf{v}_i + \mathbf{v}_j)$ coupled with the assumption that $\#_1(\mathbf{v})$ is even as well as the fact that $\mathbf{v}_i + \mathbf{v}_j \in \{-2, 0, 2\}$, imply that $\#_1(\mathbf{u})$ is even as well and thus $\mathbf{u} \in \mathcal{E}_n$ as required.

For the induction step, consider $\mathbf{v} \in \mathcal{E}_n$ and $\mathbf{u} \in \{-1,1\}^n$ satisfying $|\mathrm{Diff}(\mathbf{v},\mathbf{u})| = 2m + 2$ for some positive integer $m$ and assume that the claim holds true for any pair of vectors $\mathbf{x} \in \mathcal{E}_n$ and $\mathbf{y} \in \{-1,1\}^n$ satisfying $|\mathrm{Diff}(\mathbf{x},\mathbf{y})| = 2k$ for some positive integer $k \leqslant m$. Let $1 \leqslant i < j \leqslant n$ be any two distinct indices for which $\mathbf{v}_i \neq \mathbf{u}_i$ and $\mathbf{v}_j \neq \mathbf{u}_j$ both hold. The vector

$$\mathbf{v}' := (\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, -\mathbf{v}_i, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_{j-1}, -\mathbf{v}_j, \mathbf{v}_{j+1}, \ldots, \mathbf{v}_n)$$

satisfies $|\mathrm{Diff}(\mathbf{v}, \mathbf{v}')| = 2$; hence, $\mathbf{v}' \in \mathcal{E}_n$ holds by the induction hypothesis. Since, moreover, $|\mathrm{Diff}(\mathbf{u}, \mathbf{v}')| = 2m$, it follows by the induction hypothesis that $\mathbf{u} \in \mathcal{E}_n$. This concludes the proof of the lemma. □

We are now in position to prove the first main result of this section, namely Lemma 15.

*Proof of Lemma 15.* A vector $\mathbf{r} \in \{-1, 1\}^n$ is said to be *valid* if $\mathbf{r} \in \mathcal{E}_n$ and $\langle \mathbf{r}, \mathbf{x} \rangle = 2k$. Since $|\mathcal{E}_n| = 2^{n-1}$, it suffices to prove that there are $\binom{n}{n/2+k}$ valid vectors. In light of (6), it remains to prove that there is a bijection from the set of valid vectors to the set $S_k$.

Given a valid vector $\mathbf{r}$ (such a vector exists by the premise of the lemma), define $\varphi(\mathbf{r}) := (\mathbf{r}_1\mathbf{x}_1, \ldots, \mathbf{r}_n\mathbf{x}_n) \in \{-1, 1\}^n$. The validity of $\mathbf{r}$ implies that $\sum_{i=1}^n \varphi(\mathbf{r})_i = 2k$ and thus $\varphi(\mathbf{r}) \in S_k$. To see that $\varphi(\cdot)$ is injective, note that given two different valid vectors $\mathbf{r}$ and $\mathbf{r}'$, there exists an index $i \in [n]$ such that $\mathbf{r}_i \neq \mathbf{r}'_i$. As $\mathbf{x}$ is fixed, this compels that $\varphi(\mathbf{r})_i = \mathbf{r}_i\mathbf{x}_i \neq \mathbf{r}'_i\mathbf{x}_i = \varphi(\mathbf{r}')_i$ so that $\varphi(\mathbf{r}) \neq \varphi(\mathbf{r}')$.

To prove that $\varphi(\cdot)$ is surjective, fix $\mathbf{v} \in S_k$ and define the vector $\mathbf{y} \in \{-1, 1\}^n$ whose entries are uniquely determined by the equalities $\mathbf{v}_i = \mathbf{y}_i\mathbf{x}_i$, that is, for every $i \in [n]$, if $\mathbf{v}_i = \mathbf{x}_i$, then $\mathbf{y}_i = 1$, and otherwise $\mathbf{y}_i = -1$. It is evident that, if $\mathbf{y}$ is valid, then $\mathbf{v} = \varphi(\mathbf{y})$. Since, moreover, $\mathbf{v} \in S_k$, it suffices to prove that $\mathbf{y} \in \mathcal{E}_n$. To that end, let $\mathbf{r}$ be an arbitrary valid vector. Since $\sum_{i=1}^n \varphi(\mathbf{r})_i = 2k = \sum_{i=1}^n \mathbf{v}_i$, it follows by Lemma 31 that $|\mathrm{Diff}(\mathbf{v}, \varphi(\mathbf{r}))|$ is even. Note that $\mathbf{y}_i = \mathbf{r}_i$ whenever $i \notin \mathrm{Diff}(\mathbf{v}, \varphi(\mathbf{r}))$, and $\mathbf{y}_i = -\mathbf{r}_i$ whenever $i \in \mathrm{Diff}(\mathbf{v}, \varphi(\mathbf{r}))$. Consequently, $|\mathrm{Diff}(\mathbf{y}, \mathbf{r})|$ is even and thus $\mathbf{y}$ is even by Lemma 32. □

We conclude this section with a proof of Lemma 16.

*Proof of Lemma 16.* Since $\#_1(\mathbf{x}) \equiv \#_1(\mathbf{y}) \pmod 2$ holds by assumption, it follows by Observation 30 that $|\mathrm{Diff}(\mathbf{x}, \mathbf{y})| = 2m$ for some non-negative integer $m$. The set $\mathrm{Diff}(\mathbf{x}, \mathbf{y})$ having even cardinality has two useful implications. The first is that $n - |\mathrm{Diff}(\mathbf{x}, \mathbf{y})|$ is an even integer; this on account of $n$ being even by assumption. Using the previously introduced notation $\alpha n := \alpha(\mathbf{x}, \mathbf{y})n := n - |\mathrm{Diff}(\mathbf{x}, \mathbf{y})|$, we infer that $\alpha n$ and $(1 - \alpha)n$ are both even integers.

The second implication is that $\langle \mathbf{v}, \mathbf{x} \rangle = \langle \mathbf{v}, \mathbf{y} \rangle + \ell$, for some $\ell \in \{4k : k \in \mathbb{Z}, -m \leqslant k \leqslant m\}$, holds for every $\mathbf{v} \in \{-1, 1\}^n$. Indeed, reaching $\langle \mathbf{v}, \mathbf{x} \rangle$ starting from $\langle \mathbf{v}, \mathbf{y} \rangle$ entails iterating over each member of the even-sized set $\mathrm{Diff}(\mathbf{x}, \mathbf{y})$ and adding or subtracting two from the current value accumulated thus far.

If, additionally, $\langle \mathbf{v}, \mathbf{x} \rangle = 2k_{\mathbf{x}}$ and $\langle \mathbf{v}, \mathbf{y} \rangle = 2k_{\mathbf{y}}$, where $k_{\mathbf{x}}$ and $k_{\mathbf{y}}$ are integers, then $k_{\mathbf{x}} \equiv k_{\mathbf{y}} \pmod 2$, for indeed

$$k_{\mathbf{x}} - k_{\mathbf{y}} = \frac{\langle \mathbf{v}, \mathbf{x} \rangle - \langle \mathbf{v}, \mathbf{y} \rangle}{2} = \frac{\ell}{2} \in 2\mathbb{Z}.$$

Given $\mathbf{v} \in \{-1, 1\}^n$, set

$$S_1(\mathbf{v}) := \{i \in [n] \setminus \mathrm{Diff}(\mathbf{x}, \mathbf{y}) : \mathbf{v}_i\mathbf{x}_i = 1\} \text{ and } S_2(\mathbf{v}) := \{i \in \mathrm{Diff}(\mathbf{x}, \mathbf{y}) : \mathbf{v}_i\mathbf{x}_i = 1\}.$$

Additionally, set

$$\bar{S}_1(\mathbf{v}) := \Big([n] \setminus \mathrm{Diff}(\mathbf{x}, \mathbf{y})\Big) \setminus S_1(\mathbf{v}) \text{ and } \bar{S}_2(\mathbf{v}) := \mathrm{Diff}(\mathbf{x}, \mathbf{y}) \setminus S_2(\mathbf{v}).$$

There exist integers $m_1 := m_1(\mathbf{v})$ and $m_2 := m_2(\mathbf{v})$ such that $|S_1(\mathbf{v})| = \frac{\alpha n}{2} + m_1$ and $|S_2(\mathbf{v})| = \frac{(1-\alpha)n}{2} + m_2$. If $\langle \mathbf{v}, \mathbf{x} \rangle = 2k_{\mathbf{x}}$ for some integer $k_{\mathbf{x}}$, then

$$2k_{\mathbf{x}} = \sum_{i \in S_1(\mathbf{v})} 1 + \sum_{i \in \bar{S}_1(\mathbf{v})} (-1) + \sum_{i \in S_2(\mathbf{v})} 1 + \sum_{i \in \bar{S}_2(\mathbf{v})} (-1) = 2m_1 + 2m_2.$$

Using the definition of $\mathrm{Diff}(\mathbf{x}, \mathbf{y})$, an analogous argument shows that if $\langle \mathbf{v}, \mathbf{y} \rangle = 2k_{\mathbf{y}}$ for some integer $k_{\mathbf{y}}$, then $2k_{\mathbf{y}} = 2m_1 - 2m_2$.

Therefore[12]

$$m_1 = \frac{k_{\mathbf{x}} + k_{\mathbf{y}}}{2} \quad \text{and} \quad m_2 = \frac{k_{\mathbf{x}} - k_{\mathbf{y}}}{2};$$

in particular, $m_1$ and $m_2$ are independent of $\mathbf{v}$. We conclude that the number of vectors $\mathbf{r} \in \mathcal{E}_n$ for which $\langle \mathbf{r}, \mathbf{x} \rangle = 2k_{\mathbf{x}}$ and $\langle \mathbf{r}, \mathbf{y} \rangle = 2k_{\mathbf{y}}$ both hold is

$$\binom{\alpha n}{\frac{\alpha n}{2} + m_1} \binom{(1-\alpha)n}{\frac{(1-\alpha)n}{2} + m_2}.$$

Since, moreover, $|\mathcal{E}_n| = 2^{n-1}$, it follows that

$$\mathbb{P}\Big[\langle \mathbf{r}, \mathbf{x} \rangle = 2k_{\mathbf{x}}, \langle \mathbf{r}, \mathbf{y} \rangle = 2k_{\mathbf{y}}\Big] = \frac{1}{2^{n-1}} \binom{\alpha n}{\frac{\alpha n + k_{\mathbf{x}} + k_{\mathbf{y}}}{2}} \binom{(1-\alpha)n}{\frac{(1-\alpha)n + k_{\mathbf{x}} - k_{\mathbf{y}}}{2}}$$

as claimed. □

[12]Recall that $k_{\mathbf{x}} \equiv k_{\mathbf{y}} \pmod 2$ so that $k_{\mathbf{x}} \pm k_{\mathbf{y}}$ is even.