The intransitive dice kernel: $\frac{\mathbb{1}_{x \geq y} - \mathbb{1}_{x \leq y}}{4} - \frac{3(x-y)(1+xy)}{8}$

Ashwin Sah[1] · Mehtaab Sawhney[1]

Abstract

Answering a pair of questions of Conrey, Gabbard, Grant, Liu, and Morrison, we prove that a triplet of dice drawn from the *multiset model* are intransitive with probability $1/4 + o(1)$ and the probability a random pair of dice tie tends toward αn^{-1} for an explicitly defined constant α. This extends and sharpens the recent results of Polymath regarding the *balanced sequence model*. We further show the distribution of larger tournaments converges to a universal tournamenton in both models. This limit naturally arises from the discrete spectrum of a certain skew-symmetric operator (given by the kernel in the title acting on $L^2([-1, 1])$). The limit exhibits a degree of symmetry and can be used to prove that, for instance, the limiting probability that A_i beats A_{i+1} for $1 \leq i \leq 4$ and that A_5 beats A_1 is $1/32 + o(1)$. Furthermore, the limiting tournamenton has range contained in the discrete set $\{0, 1\}$. This proves that the associated tournamenton is non-quasirandom in a dramatic fashion, vastly extending work of Cornacchia and Hązła regarding the continuous analogue of the balanced sequence model. The proof is based on a reduction to conditional central limit theorems (related to work of Polymath), the use of a "Poissonization" style method to reduce to computations with independent random variables, and the systematic use of switching-based arguments to extract cancellations in Fourier estimates when establishing local limit-type estimates.

Mathematics Subject Classification 60F05 · 60C05

✉ Ashwin Sah
asah@mit.edu

Mehtaab Sawhney
msawhney@mit.edu

[1] Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

1 Introduction

We consider the following pair of models of random dice.

Definition 1.1 A n-sided die is a sequence of numbers $(a_1, \ldots, a_n) \in [n]^n$ such that $\sum_{j=1}^n a_j = n(n+1)/2$. In the *multiset model*, the faces of a die $(a_1, \ldots, a_n)$ are sampled as a uniform random nondecreasing sequence in $[n]$ which satisfy $\sum_{j=1}^n a_j = n(n+1)/2$. In the *balanced sequence model* the faces of a die $(a_1, \ldots, a_n)$ are sampled as a uniform random sequence in $[n]$ such that $\sum_{j=1}^n a_j = n(n+1)/2$.

We also require a notion of when one die is said to "beat" another die.

Definition 1.2 An n-sided die $(a_1, \ldots, a_n)$ *beats* another die $(b_1, \ldots, b_n)$ if

$$\sum_{j=1}^n \sum_{k=1}^n \left(\mathbb{1}_{a_j > b_k} + \frac{1}{2} \mathbb{1}_{a_j = b_k} \right) > \frac{n^2}{2}.$$

Furthermore we say that die $(a_1, \ldots, a_n)$ *ties* die $(b_1, \ldots, b_n)$ if

$$\sum_{j=1}^n \sum_{k=1}^n \left(\mathbb{1}_{a_j > b_k} + \frac{1}{2} \mathbb{1}_{a_j = b_k} \right) = \frac{n^2}{2}.$$

Our goal is to study *dice tournaments*. Specifically, we sample m independent random n-sided dice, either all from the multiset model or all from the balanced sequence model, and consider the outcome of each pair. We will think of m as fixed while n is tending to infinity.

The phenomenon of intransitive dice are exemplified by an example constructed by Efron in the 1960s [11]: consider the dice[1]

$$A = (0, 0, 4, 4, 4, 4), \ B = (3, 3, 3, 3, 3, 3), \ C = (2, 2, 2, 2, 6, 6), \ D = (1, 1, 1, 5, 5, 5).$$

Efron observed that in this example that A beats B, B beats C, C beats D, and D beats A: peculiarly, the relation "beats" is not transitive. This phenomenon gathered a substantial amount of popular interest [12, 23] including appearing in Martin Gardner's column in Scientific American [11].

Mathematical work until recently had largely been focused on constructing tournaments with various properties [1–3, 10, 18, 28]; for instance work of Moon and Moser [18] established that given any tournament T there exists a set of dice (not necessarily satisfying the sum constraints of Definition 1.1) which realize this tournament T.[2] However, recently there has been significant interest in understanding random models of intransitive dice due to a set of conjectures raised in work of Conrey et al. [7].

---

[1] Notice that as stated these dice do not satisfy the precise sum and face side bounds specified in Definition 1.1.

[2] As it turns out, our main results on random intransitive dice can be used to reprove a number of these results; we refer the reader to Proposition 8.3.

In the work of Conrey et al. [7], the authors considered dice drawn from the multiset model. While a nice model for dice, one may ask why they do not consider the "most natural" model of dice where there is no additional condition on the sum. In this case it is straightforward to observe empirically (and can be proven rigorously) that with high probability whether die A beats die B can be determined simply by looking at the sum of the faces of the dice. Conrey et al. [7] conducted empirical simulations in the multiset model. Based on these experimental results, they conjectured ([7, Conjectures 1,2,3]) that as $n \to \infty$ (a) the probability a pair of dice tie is $o(1)$ (b) for a random triplet of dice A, B, and C the probability that A beats B, B beats C, and C beats A is $1/8 + o(1)$ and (c) the tournament associated to dice is quasirandom. (Conrey et al. [7] equivalently formulate (c) in terms of the probability of various m-die tournaments.)

The first rigorous progress towards these conjectures was made by Polymath [22], where they considered n-sided die drawn from not the multiset model but from the balanced-sequence model in Definition 1.1. In this balanced sequence model, Polymath [22] was able to prove both conjectures (a) and (b) by showing that for almost all dice A drawn from the balanced sequence model, approximately half of the dice from the balanced sequence model beat it. However, based on numerical calculations Polymath conjectured that (c) is false (see discussion surrounding [22, Conjecture 1.3]). This suspicion was later confirmed in a continuous analogue of the balanced sequence model by work of Cornacchia and Hązła [8] where die faces are sampled from [0, 1] uniformly at random. They proved this by studying four-cycle counts and proved that there exists a small absolute constant $\varepsilon > 0$ such that the probability that A beats B, B beats C, C beats D, and D beats A for n large is at least $1/16 + \varepsilon$ (higher than if the underlying tournament was quasirandom). Finally, in the work of Hązła et al. [14], the phenomenon of transitivity was investigated in the context of die faces which are drawn independently at random from a fixed distribution ρ which is continuous. Remarkably, the phenomenon of intransitivity is extremely delicate and under mild conditions on ρ the only distribution exhibiting any form of intransitivity is the uniform distribution. In all the rigorous work regarding probabilistic models of intransitive dice, the use of local central limit theorem type techniques has been crucial and this has been aided by the fact that the underlying faces of the die are independent modulo conditioning on a simple linear relation. We note this is no longer true in the original multiset model of Conrey et al. [7] and this served as a key obstacle for extending results to the original model.

Our main result is a complete characterization of the tournament associated with intransitive dice. Our results are sufficiently strong to naturally explain the results of Polymath [22] and Cornacchia and Hązła [8] and point to a number of surprising phenomena which are not immediately obvious numerically.

In order to state our main result we will require the definition of a certain operator on $L^2([-1, 1])$.

Definition 1.3 Consider the skew-symmetric kernel $f : [-1, 1]^2 \to \mathbb{R}$ defined by

$$f(x, y) = \frac{\mathbb{1}_{x \geq y} - \mathbb{1}_{x \leq y}}{4} - \frac{3(x - y)(1 + xy)}{8}.$$

Define the operator $\mathcal{A} \colon L^2([-1, 1]) \to L^2([-1, 1])$ (with the Lebesgue measure) by

$$\mathcal{A}(g)(x) = \int_{-1}^{1} f(x, y)g(y)dy.$$

Let $\sigma_1 \geq \sigma_2 \geq \cdots$ denote real numbers so that $\{\pm i\sigma_\ell \colon \ell \geq 1\}$ forms the discrete spectrum of $\mathcal{A}$.

Remark Since $\mathcal{A}$ is real skew-symmetric we have that the spectrum is purely imaginary and coming in pairs. Furthermore, based on numerical computation, a closed form solution for σ_j appears unlikely.

Our main result captures the precise probability distribution associated with the dice tournament.

Theorem 1.4 *Fix $m \geq 2$ and independently sample n-sided dice $A_1, \ldots, A_m$, either all from the multiset model or all from the balanced sequence model. Let $G^{(j)}$ for $1 \leq j \leq m$ be infinite vectors of standard Gaussians and for $1 \leq j < k \leq m$ let*

$$H_{jk} = \sum_{\ell \geq 1} \sigma_\ell \left(G_{2\ell-1}^{(j)} G_{2\ell}^{(k)} - G_{2\ell}^{(j)} G_{2\ell-1}^{(k)} \right).$$

Then for any digraph D on vertices $[m]$,

$$\lim_{n \to \infty} \mathbb{P}[A_j \text{ beats } A_k \text{ for all } jk \in E(D)] = \mathbb{P}[H_{jk} > 0 \text{ for all } jk \in E(D)].$$

Remark H_{jk} is defined by a convergent sum almost surely due to the bound $\sum_{\ell \geq t} \sigma_\ell^2 = O(1/t)$, which we prove in Lemma 2.6 (**M9**), and an application of Borel–Cantelli to the random events $\mathcal{E}_t$ defined by $|\sum_{t \leq \ell < 2t} \sigma_\ell (G_{2\ell-1}^{(j)} G_{2\ell}^{(k)} - G_{2\ell}^{(j)} G_{2\ell-1}^{(k)})| \geq t^{-1/4}$ for t ranging over powers of 2. Indeed, $\mathbb{P}[\mathcal{E}_t] = O(t^{-1/2})$ by the Chebyshev inequality, which has finite sum over powers of 2, so all but finitely many $\mathcal{E}_t$ hold and the convergence follows. Alternatively, we can interpret each individual H_{jk} as a Gaussian with random variance equal to the inverse of an almost surely convergent weighted sum of chi-squared distributions.

Remark The proof of Theorem 1.4 actually shows something stronger, which is that $(H_{jk})_{1 \leq j < k \leq m}$ is the limiting distribution of $(c \cdot \text{margin}_{jk}/n)_{1 \leq j < k \leq m}$, where $c = 1/2$ for the multiset model and $c = 1$ for the balanced sequence model and where margin_{jk} is by how much die A_j beats A_k (i.e., how many more pairs than $n^2/2$, possibly negative, A_j beats A_k for).

We note that the statement of Theorem 1.4 may appear slightly strange and difficult to work with; however, a number combinatorial consequences follow in a routine manner given Theorem 1.4.

Corollary 1.5 *Sample m independent random n-sided dice $A_1, \ldots, A_m$ either all from the multiset model or all from the balanced sequence model. Then for any digraph*

D on vertices [m] let D_v denote the digraph where all edges incident to the vertex $v \in [m]$ are reversed. We have

$$lim_{n \to \infty} \mathbb{P}[A_j \text{ beats } A_k \text{ for all } jk \in E(D)] = \lim_{n \to \infty} \mathbb{P}[A_j \text{ beats } A_k \text{ for all } jk \in E(D_v)]$$

for all $v \in [m]$. Furthermore let D' denote the digraph where all the edges of D are reversed. We have

$$\mathbb{P}[A_j \text{ beats } A_k \text{ for all } jk \in E(D)] = \mathbb{P}[A_j \text{ beats } A_k \text{ for all } jk \in E(D')].$$

From Theorem 1.4 we see that the probability a pair of dice tie is $o(1)$. Then, considering D to be a directed cycle on 3 vertices and comparing to D_v, along with using permutation symmetry, we immediately see that all labelled 3-vertex tournaments appear asymptotically with the same probability. Thus a random triplet of dice is intransitive with probability $1/4 + o(1)$. This immediately implies the conjectures of Conrey et al. [7, Conjectures 1, 2] (and recovers the results of Polymath which proved these two facts in the balanced sequence model).

We can deduce that a forest with e edges occurs with probability $2^{-e} + o(1)$ by iteratively applying Corollary 1.5 to a leaf (and using that ties occur negligibly). We can also deduce that any orientation of a labeled $(2k + 1)$-cycle occurs with the same limiting probability $2^{-(2k+1)} + o(1)$ by repeatedly applying the two operations specified in Corollary 1.5. These are perhaps surprising given the results of Cornacchia and Hązła [8] showing a lack of quasirandomness for continuous dice models. We conjecture, however, that the only equalities between complete tournaments in the limit can be achieved via these symmetries and permutation symmetry.

For our next corollary, we will require the tournament analogue of a graphon. We refer the reader to [29, Chapter 4] for a more extensive discussion of graphons.

Definition 1.6 Given two measurable functions $U, W : [0, 1]^2 \to \mathbb{R}$, define the *cut metric* as

$$\delta_\square(U, W) = \inf_\phi \sup_{S,T \in [0,1]} \left| \int_{S \times T} U(x, y) - W(\phi(x), \phi(y)) dx dy \right|,$$

where the infimum ϕ is taken over all invertible measure preserving maps. We define the *tournamentons* T_0 to be the space of all functions $T : [0, 1]^2 \to [0, 1]$ such that $T(x, y) = 1 - T(y, x)$ and let $\widetilde{T_0}$ denote the space of tournamentons modulo identifying tournamentons with cut distance 0.

As is standard one can identify a graph G with an associated graphon, and similar for a tournamenton, by embedding the adjacency matrix into $[0, 1]^2$ (for the tournamenton this requires putting values of $1/2$ on the diagonal); we will carry this transformation out without comment.

Corollary 1.7 *Consider the graph T_n where the vertex set is either (a) all nondecreasing sequences $(a_1, \ldots, a_n)$ in $[n]^n$ such that $\sum_{j=1}^n a_j = n(n + 1)/2$ or (b) all sequences*

$(a_1, \ldots, a_n)$ in $[n]^n$ such that $\sum_{j=1}^n a_j = n(n+1)/2$, and where there is a directed edge from one sequence to another if the corresponding die beats the other.

Then T_n converges under the cut metric to a tournamenton $\mathcal{T}$ (which is the same in cases (a) and (b)). Furthermore, the preimage of the set $\{0, 1\}$ under $\mathcal{T}$ has measure 1.

Remark Technically T_n may not be a tournament but a partial tournament due to ties, so *a priori* we only have convergence to a *partial tournamenton*; however, a consequence of Theorem 1.4 discussed above is that ties occur with probability $o(1)$ so we will obtain a genuine tournamenton in the limit.

Note that the density of digraph D in the tournament T_n is precisely the probability that the associated digraph of dice beating other dice occurs when sampling from either the multiset model (case (a)) or the balanced sequence model (case (b)). Thus the density of digraph D in the limit tournament $\mathcal{T}$ is the limiting probability described by Theorem 1.4.

The claim that the preimage of the set $\{0, 1\}$ has measure 1 is equivalent to the fact that for every $\varepsilon > 0$ there is a k such that a $\mathcal{T}$-random tournament (defined analogously to a W-random graph [29, Section 4.4]) on k vertices lies in a set of size $2^{\varepsilon k^2}$ with at probability at least $1 - \varepsilon$. This equivalence is detailed in Lemma 8.1; we will prove Corollary 1.7 through this equivalence and prove that one can take a polynomial relation between k and ε. The fact that $\mathcal{T} \neq 1/2$ corresponds to a lack of quasirandomness. We also establish that the directed 4-cycle in particular occurs with limiting probability greater than $1/16$ in Proposition 8.2, and show that all digraphs D have positive density in $\mathcal{T}$ in Proposition 8.3.

Finally, we also precisely quantify the probability that a given pair of dice are tied beyond the $o(1)$ guaranteed as a consequence of Theorem 1.4.

Theorem 1.8 *Let A and B be dice which are jointly drawn independently from the multiset model. Let $\alpha = 2^{-5/2}\pi^{-1/2}\mathbb{E}[(\sum_{\ell \geq 1} \sigma_\ell^2(Z_\ell^2 + Z_\ell'^2))^{-1/2}]$, where $Z_\ell, Z_\ell' \sim \mathcal{N}(0, 1)$. We have*

$$\mathbb{P}[A \text{ ties } B] = (\alpha + o(1))n^{-1}$$

for some absolute constant $c = c_{1.8} > 0$. If instead A and B are jointly drawn independently from the balanced sequence model then

$$\mathbb{P}[A \text{ ties } B] = (2\alpha + o(1))n^{-1}.$$

Remark This can be heuristically reconstructed by considering the second remark after Theorem 1.4 with $m = 2$. H_{12} is the limiting distribution of $c \cdot \text{margin}_{12}/n$ (where $c = 1/2$ for the multiset model and $c = 1$ for the balanced sequence model). If we imagine that the mass of this distribution was discretized in the obvious way along all possible values of margin_{12} in the lattice $\mathbb{Z}/2$, we obtain the above. In fact, one can use the techniques in Sect. 9 to show a local limit theorem for margin_{12}:

$$\mathbb{P}[\text{margin}_{12} = x] = \frac{c}{2n} f_{H_{12}}(cx/n) + o(1/n)$$

uniformly for $x \in \mathbb{Z}/2$, where $f_{H_{12}}$ is the probability density function of H_{12}. We do not prove this here since the technical details are quite involved, but note that Theorem 1.8 is the $x = 0$ case.

We interpret the constant α as the (inverse) standard deviation around the best linear approximant (in the sense of Ordinary Least Squares) to a conditioned Brownian motion at the end of Sect. 9.

In general, a tournament T with exactly t ties among m dice, and $\binom{m}{2} - t$ prescribed outcomes of the other match-ups, where m and $0 \le t \le \binom{m}{2}$ are fixed, should occur with probability $(c_T + o(1))n^{-t}$. We do not pursue such a general statement here though similar techniques may apply and a probabilistic interpretation of the constant c_T should arise from Theorem 1.4 similar to the case $(m, t) = (2, 1)$ above.

1.1 First steps, proof outline, and organization

Our techniques at a high level involve Fourier analysis in the style of local limit theorems. In particular, we study various "conditional Fourier coefficients" in detail to show that the normalized joint distribution of "victory margins" (see the second remark following Theorem 1.4) converges to $(H_{jk})_{1 \le j < k \le m}$. We also use a more detailed analysis involving additional control on the "coarseness" of certain modified statistics of random dice to get very good local control of the event that there is a precise tie. We defer a more detailed proof outline to Sect. 3 after developing the basic tools to attack the problem in Sect. 2.

The first step in proving Theorem 1.4 (in the multiset model) relies on observing that while the dice face in the multiset model are nonindependent, the frequency statistics can be given a natural "near-independent" model. This ultimately relies on a well-known bijection between the multiset model and the simple random walk; the details appear in Lemma 2.2. We note that in the context of the balanced sequence model, Lemma 2.2 reduces to the "Poissonization" trick. Given this we interpret the "beats" relation through frequency counts (Lemmas 2.3, 2.5) and the operator in Definition 1.3 arises naturally. These initial steps are carried out in Sect. 2, and provide the key starting point to understand the necessary distributions from a Fourier perspective.

Given the setup in Sect. 2, we provide a heuristic outline of the argument for Theorems 1.4 and 1.8 and an overview of the various consequences in Sect. 3. We then collect a list of technical preliminaries which will be used throughout the paper in Sect. 4. We prove various Fourier coefficient bounds used in the proofs of Theorems 1.4 and 1.8 in Sect. 5. We prove Theorem 1.4 in Sect. 6, modulo a technical ingredient proven in Sect. 7, and then collect various consequences following from Theorem 1.4 in Sect. 8. Finally we prove Theorem 1.8 in Sect. 9.

Notation

We write $f = O(g)$ to mean that $f \le Cg$ for some absolute constant C, and $g = \Omega(f)$ and $f \lesssim g$ to mean the same. We write $f = o(g)$ if for all $c > 0$ we have $f \le cg$ once the implicit growing parameter (typically n) grows large enough, and $g = \omega(f)$ means

the same. We also use $\asymp$ to denote two quantities which are within absolute constants. Subscripts imply a dependence of these implicit constants on those parameters. We use $\overset{d.}{=}, \overset{d.}{\to}$ for distributional equality and limits, respectively.

For $\mu \in \mathbb{R}^d$ and positive semidefinite $\Sigma \in \mathbb{R}^{d \times d}$ we let $\mathcal{N}(\mu, \Sigma)$ be the Gaussian vector with mean μ and covariance matrix Σ. For finite matrices M we will use M_{ij} to denote the entry in the (i, j) position. Throughout this paper all logarithms are base e.

2 Count statistics of balanced sequence model and multiset model

The idea to get a handle on the multiset model is to create a procedure for sampling which derives from a sequence of independent random variables. We will require the notion of a frequency statistic which will be crucial for our purposes.

Definition 2.1 Given an n-sided die $A = (a_1, \ldots, a_n)$ define the *frequency counts* of A to be

$$\widetilde{a}_i = |\{j : a_j = i\}|$$

for $1 \le i \le n$.

The key point is the following distributional claim regarding the frequency counts of a die drawn from either multiset or balanced sequence model, which relates these models to a sequence of either geometric (in the multiset case) or a sequence of Poisson random variables (in the balanced sequence case). In the balanced sequence case this is essentially equivalent to the "Poissonization" trick.

Lemma 2.2 *We have the following:*

- *If B is drawn from the multiset model we have*

$$(\widetilde{b}_1, \ldots, \widetilde{b}_n) \overset{d.}{=} (G_1, \ldots, G_n),$$

  where G_j are sampled as follows: draw independent $\mathrm{Geom}(1/2)$ random variables[3] G_i and then condition on $\sum_{j=1}^n G_j = n$ and $\sum_{j=1}^n j G_j = n(n+1)/2$.
- *If B is drawn from the balanced sequence model we have*

$$(\widetilde{b}_1, \ldots, \widetilde{b}_n) \overset{d.}{=} (P_1, \ldots, P_n),$$

  where P_j are sampled as follows: draw independent $\mathrm{Pois}(1)$ random variables P_j and then condition on $\sum_{j=1}^n P_j = n$ and $\sum_{j=1}^n j P_j = n(n+1)/2$.

---

[3] Here $X \overset{d.}{=} \mathrm{Geom}(1/2)$ means $\mathbb{P}[X = k] = (1/2)^{k+1}$ for $k \in \{0, 1, \ldots\}$. Note this is 0-indexed, corresponding to the number of "failures" before a repeatedly flipped fair coin shows heads, instead of the number of "trials".

Proof We consider the first case. Notice that the multiset model of a die can equivalently be sampled by sampling a uniformly random right-up walk between $(1, 1)$ and $(n + 1, n)$ and looking at the height of each rightward step, conditional on the area under the walk being $n(n + 1)/2$. Indeed, there is a standard bijection between nondecreasing integer sequences $(b_1, \ldots, b_n)$ with $1 \leq b_j \leq n$ and such walks: each a_j corresponds to a rightward step from (j, b_j) to $(j + 1, b_j)$; furthermore, the area under the walk ends up being $b_1 + \cdots + b_n$. Notice that drawing such a walk is equivalent to looking at an infinite random walk which takes steps in the directions $(1, 0)$ or $(0, 1)$ each with probability $1/2$ and then conditioning on starting at $(1, 1)$ and passing through both $(n + 1, n)$ and $(n + 1, n + 1)$, then truncating appropriately.

Now define G_j as precisely the length of the horizontal segment on the line $y = j$ in this conditioned infinite random walk. In the unconditioned infinite walk starting at $(1, 1)$, we have that the lengths (which might be 0) of these horizontal segments in order have an independent distribution where the law is by definition a sequence of independent geometric random variables with parameter $1/2$. Notice that conditioning on the walk passing through the line segment $(n+1, n)$ and $(n+1, n+1)$ is equivalent to $\sum_{j=1}^{n} G_j = n$ and the conditioning on area is equivalent to $\sum_{j=1}^{n} j G_j = n(n+1)/2$. Considering the bijection defined above, these conditioned G_j then correspond directly to the $\widetilde{b}_j$.

The second case is simpler. Notice that if one draws n faces from $[n]$ uniformly at random then we have the proportionality

$$\mathbb{P}[(\widetilde{b}_1, \ldots, \widetilde{b}_n) = (c_1, \ldots, c_n)] \propto \prod_{j=1}^{n} \frac{1}{c_j!}$$

for tuples $(c_j)_{1 \leq j \leq n} \in \{0, \ldots, n\}^n$ with sum n. The result then follows since $\mathbb{P}[\text{Pois}(1) = k] = \frac{e^{-1}}{k!}$ for $k \in \mathbb{Z}$, and since conditioning on the sum of the dice being $n(n + 1)/2$ corresponds to conditioning on $\sum_{j=1}^{n} j P_j = n(n+1)/2$. $\qquad\square$

The precise reason this description is useful is that given a die B one can define a linear function of the frequency count statistics of another die A which captures precisely whether A beats B or not. An equivalent computation appears in the work of Polymath [22, Section 4]; the formulation presented there however is more naturally a linear function of the "die faces" instead of the "frequency count statistics".

Lemma 2.3 *We have that a die A with sides $(a_1, \ldots, a_n)$ beats a die B with sides $(b_1, \ldots, b_n)$ if and only if*

$$\sum_{j=1}^{n} \left(\sum_{1 \leq k < j} \widetilde{b}_k + \frac{\widetilde{b}_j}{2} - (j - 1/2) \right) \widetilde{a}_j > 0$$

and ties if and only if the sum on the left is 0.

Proof Notice that

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\mathbb{1}_{a_i>b_j}+\frac{1}{2}\mathbb{1}_{a_i=b_j}\right)=\sum_{j=1}^{n}\tilde{a}_j\sum_{1\le k<j}\tilde{b}_j+\sum_{j=1}^{n}\frac{\tilde{a}_j\tilde{b}_j}{2}$$

$$=\sum_{j=1}^{n}\tilde{a}_j\left(\sum_{1\le k<j}\tilde{b}_j+\frac{\tilde{b}_j}{2}\right).$$

Since A is an n-sided die with sum of faces $n(n+1)/2$ we have

$$\sum_{j=1}^{n}\tilde{a}_j(j-1/2)=\sum_{j=1}^{n}a_j-\frac{n}{2}=\frac{n^2}{2}$$

and therefore the event that A beats B is precisely equivalent to

$$\sum_{j=1}^{n}\tilde{a}_j\left(\sum_{1\le k<j}\tilde{b}_k+\frac{\tilde{b}_j}{2}-(j-1/2)\right)>0,$$

whereas A and B being tied corresponds to the left side being equal to 0. □

We cast this condition in an equivalent form which will be useful for computations involving Gaussians.

Definition 2.4 Let I_n be the $n \times n$ identity matrix. Let $\vec{v}_1, \vec{v}_2 \in \mathbb{R}^n$ be defined by $v_{1i} = 1/\sqrt{n}$ for $1 \le i \le n$ and $v_{2i} = (i - (n+1)/2)/\sqrt{n(n^2-1)/12}$ for $1 \le i \le n$. Note these are orthogonal unit vectors. Let $M_n \in \mathbb{R}^{n\times n}$ be defined via $(M_n)_{ij} = \mathbb{1}_{i<j} + (\mathbb{1}_{i=j}/2)$ and $M_n^* \in \mathbb{R}^{n\times n}$ via

$$M_n^* = (I_n - \vec{v}_2\vec{v}_2^\mathsf{T})(I_n - \vec{v}_1\vec{v}_1^\mathsf{T})M_n(I_n - \vec{v}_1\vec{v}_1^\mathsf{T})(I_n - \vec{v}_2\vec{v}_2^\mathsf{T})$$

$$= (I_n - \vec{v}_1\vec{v}_1^\mathsf{T} - \vec{v}_2\vec{v}_2^\mathsf{T})M_n(I_n - \vec{v}_1\vec{v}_1^\mathsf{T} - \vec{v}_2\vec{v}_2^\mathsf{T}).$$

Equivalently, we re-express the (asymmetric) bilinear form M_n in a basis including $\vec{v}_1, \vec{v}_2$ on both sides, zero out the rows and columns corresponding to $\vec{v}_1, \vec{v}_2$, and then convert back. Finally, define $\sigma_{n,1} \ge \cdots \ge \sigma_{n,\lfloor n/2\rfloor}$ be such that $\{\pm i\sigma_{n,\ell} : \ell \in [\lfloor n/2\rfloor]\}$ is the spectrum of M_n^* (or the spectrum minus a copy of 0 if n is odd).

The following lemma introduces this discrete variant of the kernel which appears in the title of the paper.

Lemma 2.5 *Given dice A, B with frequency vectors $\tilde{a}, \tilde{b} \in \{0, \dots, n\}^n$, we have that A beats B if and only if*

$$\tilde{b}^\mathsf{T} M_n^* \tilde{a} > 0.$$

Proof This is immediate since a simple manipulation of Lemma 2.3 shows the condition that A beats B is equivalent to $(\tilde{b} - \sqrt{n}\vec{v}_1)^\mathsf{T} M_n(\tilde{a} - \sqrt{n}\vec{v}_1) > 0$, and since $\vec{v}_1^\mathsf{T}(\tilde{a} - \sqrt{n}\vec{v}_1) = \vec{v}_2^\mathsf{T}(\tilde{a} - \sqrt{n}\vec{v}_1) = \vec{v}_1^\mathsf{T}(\tilde{b} - \sqrt{n}\vec{v}_1) = \vec{v}_2^\mathsf{T}(\tilde{b} - \sqrt{n}\vec{v}_1) = 0$. □

Finally, we record some properties of M_n^* as well as $\mathcal{A}$ (Definition 1.3). We are brief with the details as it mostly amounts calculation with explicit functions and operators.

Lemma 2.6 *There exists $C = C_{2.6} > 0$ such that the following holds. Let M_n^* be as in Definition 2.4, $x = (n + 1 - 2i)/(n - 1)$ and $y = (n + 1 - 2j)/(n - 1)$. Then we have the following:*

M1

$$(M_n^*)_{ij} = \frac{\mathbb{1}_{x \geq y} - \mathbb{1}_{x \leq y}}{2} - \frac{3(x - y)(1 - 1/n)}{4} - \frac{3xy(x - y)(n - 1)^2}{4n(n + 1)}.$$

M2 *M_n^* is skew-symmetric.*
M3 *$\|M_n^*\|_{1 \to \infty} \leq C_{2.6}$ (i.e., the entries are of bounded size).*
M4 *$\|M_n^*\|_{1 \to 2} = \|M_n^{*\mathsf{T}}\|_{1 \to 2} \leq C_{2.6}\sqrt{n}$ (i.e., the row and column L^2-norms are $O(\sqrt{n})$ in size).*
M5 *$\|M_n^*\|_F/n \in [C_{2.6}^{-1}, C_{2.6}]$.[4]*
M6 *For fixed $t \geq 1$ and n sufficiently large,*

$$\sum_{\ell \geq t} \sigma_{n,\ell}^2 \leq C_{2.6} n^2/t.$$

M7 *For $1 \leq i, j, k \leq n$ we have $|(M_n^*)_{ji} - (M_n^*)_{ki}| \leq C_{2.6}|j - k|/n$ for all $i \notin [j, k] \cup [k, j]$.*

We also have the following properties of $\mathcal{A}$.

M8 *$(\sigma_{n,\ell}/n)_{1 \leq \ell \leq t} \to (\sigma_\ell)_{1 \leq \ell \leq t}$ as $n \to \infty$.*
M9 *For all $t \geq 1$ we have that $t\sigma_t \in [C_{2.6}^{-1}, C_{2.6}]$ (i.e., $t\sigma_t$ is bounded above and below by an absolute constant).*

Proof Via a direct, albeit tedious computation, one has that if $x = (n+1-2i)/(n-1)$ and $y = (n + 1 - 2j)/(n - 1)$ then

$$(M_n^*)_{ij} = \frac{\mathbb{1}_{x \geq y} - \mathbb{1}_{x \leq y}}{2} - \frac{3(x - y)(1 - 1/n)}{4} - \frac{3xy(x - y)(n - 1)^2}{4n(n + 1)}.$$

The properties **M2** to **M5** and **M7** all follow immediately via direct inspection.

To prove **M6**, it suffices to show that there is a rank $t + 4$ (say) approximation of M_n^*, call it R_t, such that $\|M_n^* - R_t\|_F^2 \lesssim n^2/t$. This follows from considering a rank t approximation for M_n and then plugging it into Definition 2.4. An appropriate rank t approximation for M_n with square-error $O(n^2/t)$ can be formed by removing square

---

[4] Here the *Frobenius norm* of matrix $M \in \mathbb{R}^{n \times n}$ is $\|M\|_F := \sqrt{\sum_{1 \leq j,k \leq n} M_{jk}^2}$.

matrices of 1s from the right isosceles triangle above the main diagonal of M_n in a dyadic fashion.

To prove the convergence given in **M8**, we proceed by an argument identifying matrices with operators in $L^2([-1, 1])$ via step functions. In particular, consider the matrices M_n^* and identify them with the kernels

$$M_n^{(*)}(x, y) = \frac{n M_n^*(\lceil n(1 - x)/2 \rceil, \lceil n(1 - y)/2 \rceil)}{2}$$

and note that the action of M_n^* on $\mathbb{R}^n$ corresponds exactly to the action of kernel $M_n^{(*)}(x, y)$ on step functions where the index $i \in [n]$ has been mapped to the interval $[1 - 2i/n, 1 - 2(i - 1)/n)$. These have the same spectrum: the multiplicative factor of $n/2$ corresponds to fact that the step function which is 1 on a single length $2/n$ interval has norm $(2/n)^{1/2}$ in the continuous formulation while it has norm 1 when viewed as a vector in $\mathbb{R}^n$.

In general, given a kernel $K : [-1, 1]^2 \to \mathbb{R}$ one can define the integral operator

$$\widetilde{K} : g(x) \to \int_{-1}^{1} K(x, y) g(y) dy$$

and we have $\|\widetilde{K}\|_{L^2([-1,1]) \to L^2([-1,1])} \leq \|K\|_{L^2([-1,1]^2)}$ by Cauchy–Schwarz (see e.g. [13, Example 9.23]). Via this identification, we have the strong convergence $M_n^{(*)}(x, y)/n \to \widetilde{M^{(*)}} := \mathcal{A}$, where the corresponding kernel is

$$f(x, y) = \frac{\mathbb{1}_{x \geq y} - \mathbb{1}_{x \leq y}}{4} - \frac{3(x - y)(1 + xy)}{8}.$$

Given this, since $\mathcal{A}, M_n^{(*)}, M_n^*$ are skew-symmetric (hence normal) operators, it is easy to see that the normalized eigenvalues of M_n^* converge to those specified by Definition 1.3 (as strong convergence implies convergence of the spectrum). This proves **M8**.

Finally we prove **M9**. In order to prove **M9**, we first note that $\mathcal{A}$ is a skew-symmetric perturbation of the integral operator associated to the function $g(x, y) = \frac{\mathbb{1}_{x \geq y} - \mathbb{1}_{x \leq y}}{4}$, with the rank of the perturbation bounded by 2. We claim that it suffices to prove that the t-th singular value of $\widetilde{g}$ scales as $\Theta(1/t)$. Indeed, apply the generalized Weyl's inequality to the Hermitian operator $\widetilde{g}^\dagger \widetilde{g}$ using that $\mathcal{A}^\dagger \mathcal{A}$ is a bounded rank perturbation.

To compute the spectrum of $\widetilde{g}$ (and thus that of $\widetilde{g}^\dagger \widetilde{g}$), note that the matrix given by $(T_n)_{ij} = \mathbb{1}_{i \geq j} - \mathbb{1}_{i \leq j}$ has characteristic polynomial $\frac{(-1)^n((\lambda+1)^n + (\lambda-1)^n)}{2}$; this is easily proven via row operations and induction. It follows that the eigenvalues of T_n are $(1 + \exp(\pi i (2j - 1)/n))/(1 - \exp(\pi i (2j - 1)/n))$ for $1 \leq j \leq n$. Thus the jth largest eigenvalue in magnitude scales as $\Theta(n/j)$. The desired result then follows by rescaling and taking $n \to \infty$. $\qquad \square$

3 Outline of the remainder of the proof

We now outline the remainder of the proofs of Theorems 1.4 and 1.8 in the multiset model; the balanced sequence model is very similar modulo adjusting various constant factors arising due to $\mathrm{Var}[\mathrm{Geom}(1/2)] = 2\mathrm{Var}[\mathrm{Pois}(1)] = 2$. We also discuss the various deductions which follow from Theorem 1.4. Consider a set of m dice $A_1, \ldots, A_m$ and let $\widetilde{a}_k = (\widetilde{a}_{kj})_{1 \le j \le n}$ be the n-dimensional vector corresponding to the frequency counts of A_k for $1 \le k \le m$.

3.1 Theorem 1.4 and its consequences

By Lemma 2.5 we have that A_j beats A_k if an only if $\widetilde{x}_k^{\mathsf{T}} M_n^* \widetilde{x}_j > 0$. Note that the constraints that $\widetilde{x}_j$ satisfy are precisely $(1, \ldots, 1)^{\mathsf{T}} \widetilde{x}_j = n$ and $(1, 2, \ldots, n)^{\mathsf{T}} \widetilde{x}_j = n(n+1)/2$ (equivalently, $\widetilde{x}_j - \vec{1}$ is orthogonal to $\vec{v}_1, \vec{v}_2$). By construction we have that $M_n^* \vec{1} = \vec{0}$ and $M_n^*(1, 2 \ldots, n) = \vec{0}$. Therefore for the sake of reasoning heuristically, we can pretend that the conditioning in Lemma 2.2 does not affect the probability distribution of $\widetilde{x}_k^{\mathsf{T}} M_n^* \widetilde{x}_j$ and instead suppose that $\widetilde{x}_\ell$ are replaced by X_ℓ, n-dimensional vectors where every entry is taken independently at random to be $\mathrm{Geom}(1/2)$. Now $X_k^{\mathsf{T}} M_n^* X_j$ is a bilinear polynomial of independent random variables. Tools such as the invariance principle of Mossel et al. [19] imply that the associated distribution is close to the distribution in the case where X_ℓ are replaced by Z_ℓ, where each entry of Z_ℓ is an independent normal of variance $\mathrm{Var}[\mathrm{Geom}(1/2)] = 2$. Given this, we can convert to a Gaussian quadratic form. This is invariant under orthogonal transformation, so a singular value decomposition for the skew-symmetric matrix M_n^* and an appropriate variant of the spectral theorem quickly leads to the distribution in Theorem 1.4. In particular, the coefficients associated in Theorem 1.4 arise precisely from an application of Lemma 2.6.

In order to prove this heuristic, we need to precisely understand the joint distribution of $(1, \ldots, 1)^{\mathsf{T}} X_j$, $(1, 2, \ldots, n)^{\mathsf{T}} X_j$, and the desired quadratic forms. We proceed using Fourier transform (characteristic function) and computing the multidimensional Fourier coefficients of the joint distribution of the quadratic forms conditional on these linear equalities. This conditional expectation can be recast using Bayes' theorem and converted to an expression involving joint coefficients involving both quadratic and linear forms, which we can provide control for using the techniques in Sect. 5. Our proof here is closely related to that of that in the work of Polymath [22] which similarly used local central limit theorem techniques to decouple various linear conditions; however the implementation is performed in a rather different manner.

We write this more explicitly. For the sake of this discussion, let $\mathcal{E}$ denote the event that all that the m samples X_ℓ for $1 \le \ell \le m$ satisfy $(1, \ldots, 1)^{\mathsf{T}} X_j = n$, $(1, 2, \ldots, n)^{\mathsf{T}} X_j = n(n+1)/2$. We then must compute

$$\mathbb{E}\left[\exp\left(i \sum_{1 \le j < k \le m} \theta_{jk} X_k^{\mathsf{T}} M_n^* X_j \right) \Big| \mathcal{E} \right]$$

for all choices of $\theta = (\theta_{jk})_{1 \leq j < k \leq m}$, where $\|\theta\|_\infty$ is roughly $\widetilde{O}(1/n)$.

Via applying Bayes' rule, this amounts to computing

$$\mathbb{E}\left[\exp\left(i \sum_{1 \leq j < k \leq m} \theta_{jk} X_k^\mathsf{T} M_n^* X_j\right) \mathbb{1}_{\mathcal{E}}\right],$$

since then considering $\theta = 0$ gives an estimate for $\mathbb{E}[\mathbb{1}_{\mathcal{E}}] = \mathbb{P}[\mathcal{E}]$ and we can divide to obtain the conditional expectation. At this juncture, much as in the work of Polymath [22], we rely on the Fourier inversion formula to convert the indicator $\mathbb{1}_{\mathcal{E}}$ into an explicit integral formula in terms of additional Fourier terms involving the above linear forms. (Note that this conversion is only available to us in the multiset model due to the key lemma Lemma 2.2, and even in the balanced sequence model we utilize the setup of Lemma 2.2 to prove Theorem 1.4.)

In particular, by applying Fourier inversion on the lattices we will find

$$\mathbb{E}\left[\exp\left(i \sum_{1 \leq j < k \leq m} \theta_{jk} X_k^\mathsf{T} M_n^* X_j\right) \mathbb{1}_{\mathcal{E}}\right]$$

$$= (2\pi)^{-2m} \int_{[-\pi,\pi]^{2m}} \mathbb{E}\left[\exp\left(i \sum_{1 \leq j < k \leq m} \theta_{jk} X_k^\mathsf{T} M_n^* X_j\right)\right.$$

$$\left. \times \exp\left(i\left(\sum_{j=1}^m \xi_{1j}\left(\sum_{r=1}^n (X_{jr} - 1)\right) + \xi_{2j}\left(\sum_{r=1}^n r(X_{jr} - 1)\right)\right)\right)\right] d\vec{\xi}.$$

In order to prove the desired result, we split the integral into several regions. If any $|\xi_{1j}| \geq n^{-1/2}(\log n)^7$ or $|\xi_{2j}| \geq n^{-3/2}(\log n)^6$, we prove that the corresponding term in the integral is super-polynomially small using Lemmas 5.6–5.8. Specifically, we condition on everything outside of the index r, and then the corresponding Fourier integral is simply a product of independent terms handled by these lemmas. To prove these lemmas, we extract cancellation in a systematic and clean manner by considering pairs and triplets of indices and performing "switches" between then in order to extract Boolean randomness. These switches allow for one to provide sufficient conditions on various coefficient sequences to be good enough to perform these arguments, and said conditions exist purely in "physical space" (whereas the approach taken in the work of Polymath [22] naturally leads one to consider how various coefficients are distributed with respect to angles on the torus). Finally, in the region where $|\xi_{1j}| \leq n^{-1/2}(\log n)^7$ and $|\xi_{2j}| \leq n^{-3/2}(\log n)^6$ we apply a Lindeberg exchange argument (see [16], and also the related proof of the invariance principle [19]) to replace the geometric random variables with Gaussians of the same variance. Using the rapid decay of Fourier coefficients the Gaussian and Gaussian rotational symmetry one can verify the Fourier coefficient matches that of the associated Gaussian prediction and thus the desired result follows via Lévy continuity and similar techniques which convert Fourier control back to physical space control.

In order to prove Corollary 1.5, we directly cite Theorem 1.4 and uses symmetries of the Gaussian distribution under negation to derive the necessary result. For Corollary 1.7, note that convergence to a tournamenton follows from general machinery since we have the convergence of each digraph. To deduce that the associated tournamenton is $\{0, 1\}$-valued we reduce to proving a random tournament on M dice takes on outcomes within a specific set of complete tournaments of size $2^{\varepsilon M^2}$ with probability at least $1 - \varepsilon$. This is shown using Theorem 1.4: note that we can simulate the limiting tournament on M vertices by sampling the Gaussians $G_\ell^{(j)}$ for $1 \le j \le M$ and $\ell \ge 1$ and computing the various H_{jk} and checking their signs. By revealing for each j the first $2M^{1/2}$ Gaussians $Z_\ell^{(j)}$ within a rounding error of M^{-10}, this provides at most $\exp(O(M^{3/2} \log M))$ buckets where almost all the probability mass lies and also allows us with good probability to determine the outcome of almost all match-ups in the tournament (this deduction requires Gaussian anticoncentration results such as Theorem 4.3 in order to see that it is unlikely that many match-ups are "too close to call" due to the rounding error). Then revealing the outcomes of the remaining match-ups introduces $\exp(o(M^2))$ total buckets that contain almost all the probability mass, and which uniquely determine the outcome of the M-die tournament.

Given the non-quasirandomness of the associated tournament from Corollary 1.7 and the underlying symmetries in Corollary 1.5 it also follows from a simple Cauchy–Schwarz argument that the limiting probability A beats B, B beats C, C beats D, and D beats A is strictly larger than $1/16$ (see Proposition 8.2). Finally, we note that via carefully choosing various Gaussians $Z_\ell^{(j)}$ to lie in certain ranges one can prove that the limiting probability of any fixed M-die tournament occurring is strictly positive (see Proposition 8.3). This allows one to quickly deduce a number of prior results as discussed in the introduction.

3.2 Proof of Theorem 1.8

To compute the probability that two dice tie, proceed via a more delicate route. As discussed in the remark following Theorem 1.8, one can see this as a (special case of a) local limit theorem version of Theorem 1.4 with two dice.

We use ideas closely related to those in the proof of Theorem 1.4, as well as additional Fourier coefficient estimates (Lemmas 5.4, 5.5) which use the extra condition that certain associated coefficient sequences "resemble a simple random walk at all scales" in a coarse sense. It follows that for almost all outcomes of die A_1, the probability a random die A_2 with frequency counts $\widetilde{x}_2$ ties A_1 is proportional to $\|M_n^* \vec{x}_2\|_2^{-1}$. (We note that such a result for the balanced sequence model is essentially implicit in the work of Polymath [22] although not stated in such a manner; however, again, our work proceeds through frequency counts instead of using independent die faces which are not available for the multiset model.)

Therefore the natural approach at this point would be to prove a limit theorem for $\|M_n^* X\|_2^2$, where X is a sequence of geometric random variables conditional on the two linear constraints $(1, \ldots, 1)^\mathsf{T} X = n$, $(1, 2, \ldots, n)^\mathsf{T} X = n(n + 1)/2$. While this appears to be possible note that $\|M_n^* \vec{x}\|_2^2$ is a genuinely quadratic polynomial in the

underlying random variables (instead of being multilinear in the case of Theorem 1.4) and hence for a direct approach various tools developed by Berkowitz [4], developed in the context of local central limit theorems for clique counts in dense random graphs, would appear to be necessary, which would greatly complicate the situation.

To circumvent this, we proceed indirectly so as to only require linear Fourier estimates. The basic idea is that given a sufficiently good upper bound on $\|M_n^* X\|_3$ (conditional on our two linear constraints), by sampling a fixed number of random coordinates $j_1, \ldots, j_T$ for some large constant T we have

$$\|M_n^* X\|_2^2 \approx \frac{n}{T} \sum_{\ell=1}^{T} \langle e_{j_\ell}, M_n^* X \rangle^2$$

holds with high probability as $T \to \infty$. Therefore the question can be reduced to a question of understanding the linear statistics $(\langle e_{j_\ell}, M_n^* X \rangle)_{1 \leq \ell \leq T}$ jointly conditional on our two linear constraints. This can be handled by precisely the techniques developed we discussed in Sect. 3.1 for Theorem 1.4. The estimates are necessarily a bit delicate since the function $y \mapsto 1/y$ is not bounded near 0 and thus care must be taken to rule out the pathological cases where $\|M_n^* X\|_2$ is small with unusually large probability.

4 Preliminaries

We briefly collect a series of preliminaries which will be used throughout the proof. First we will require a version of the classical Bernstein inequality, which generalizes Chernoff.

Theorem 4.1 ([26, Theorem 2.8.1]) *For a random variable X define the ψ_1-norm*

$$\|X\|_{\psi_1} = \inf\{t > 0 \colon \mathbb{E}[\exp(|X|/t)] \leq 2\}.$$

There is an absolute constant $c = c_{4.1} > 0$ such that the following holds. If $X_1, \ldots, X_N$ are independent random variables then

$$\mathbb{P}\left[\left|\sum_{i=1}^{N} X_i\right| \geq t\right] \leq 2 \exp\left(-c \min\left(\frac{t^2}{\sum_{i=1}^{N} \|X_i\|_{\psi_1}^2}, \frac{t}{\max_i \|X_i\|_{\psi_1}}\right)\right)$$

for all $t \geq 0$.

Next we will require the Azuma–Hoeffding inequality (see [15, Theorem 2.25]).

Lemma 4.2 (Azuma–Hoeffding inequality) *Let $X_0, \ldots, X_n$ form a martingale sequence such that $|X_k - X_{k-1}| \leq c_k$ almost surely. Then*

$$\mathbb{P}[|X_0 - X_n| \geq t] \leq 2 \exp\left(-\frac{t^2}{2 \sum_{k=1}^{n} c_k^2}\right)$$

Remark We will refer to $\sum_{k=1}^n c_k^2$ as the *variance proxy* in such a situation.

Furthermore we will require the Carbery–Wright theorem [5] for which prove that low-degree functions of Gaussians are anticoncentrated; we will only require the quadratic case.

Theorem 4.3 (see e.g. [17, Theorem 1.4]) *Fix an integer $d \geq 1$. There exists a constant C_d such that the following holds. For any $\varepsilon > 0$, if $(G_i)_{1 \leq i \leq n}$ are independent Gaussian random variables, and P is a polynomial of degree at most d then*

$$\sup_{t \in \mathbb{R}} \mathbb{P}\big[|P(G_1, \ldots, G_n) - t| \leq \varepsilon \sqrt{\mathrm{Var}(P(G_1, \ldots, G_n))}\big] \leq C_d \varepsilon^{1/d}.$$

We will also require the invariance principle of Mossel et al. [19]. The version stated in Theorem 4.5 below is a stated as [20, (11.66)] (with the necessary hypercontractivity following from [19, Proposition 3.16]).

Definition 4.4 Given a multilinear polynomial $g(x_1, \ldots, x_n) = \sum_{S \subseteq [n]} a_S \prod_{i \in S} x_i$, for $t = 1, \ldots, n$ the *influence* of the variable x_t is defined as

$$\mathrm{Inf}_t[g] = \sum_{\substack{S \subseteq [n] \\ S \ni t}} a_S^2.$$

Theorem 4.5 *Fix $M \geq 1$; there exists $M' > 0$ such that the following holds. Let g be an n-variable multilinear polynomial of degree at most k. Let $\vec{y}$ be a vector of i.i.d. random variables such that $\mathbb{E}[y_i] = 0$, $\mathbb{E}[y_i^2] = 1$ and $\mathbb{E}[|y_i|^3] \leq M$. Let $\vec{z} \sim \mathcal{N}(0, 1)^{\otimes n}$ be a vector of independent standard Gaussian random variables. Then for any three-times-differentiable function $\psi : \mathbb{R} \to \mathbb{R}$, we have*

$$\Big|\mathbb{E}[\psi(g(\vec{y})) - \psi(g(\vec{z}))]\Big| \leq (M')^k \cdot \|\psi^{(3)}\|_\infty \sum_{t=1}^n \mathrm{Inf}_t[g]^{3/2}.$$

Next, we will require the following concentration inequality for low-degree polynomials of Gaussian; the Rademacher case is stated as [20, Theorem 9.23] and the Gaussian case follows by taking limits via and applying the central limit theorem.

Theorem 4.6 *Let f be a polynomial in n variables of degree at most d. Let $\vec{x} = (x_1, \ldots, x_n)$ be a vector of independent standard Gaussian random variables. Then for any $t \geq (2e)^{d/2}$,*

$$\mathbb{P}\left[|f(\vec{x})| \geq t(\mathbb{E}[f(\vec{x})^2])^{1/2}\right] \leq \exp\left(-\frac{d}{2e} t^{2/d}\right).$$

We also require a statement allowing one to quantify the convergence in distribution of a random variable given convergence of the associated Fourier transform. The following result is immediate from [21, p. 104, Theorem 1]; this is an essentially standard inequality used in the proof of the Berry–Esseen theorem.

Theorem 4.7 *There exists an absolute constant $C = C_{4.7} > 0$ such that the following statement holds. Consider a pair of random variables X and Y and a parameter $T > 0$. We have that*

$$\sup_{\tau \in \mathbb{R}} |\mathbb{P}[X \le \tau] - \mathbb{P}[Y \le \tau]| \le C_{4.7}$$

$$\left(\int_{-T}^{T} \frac{|\mathbb{E}[\exp(itX) - \exp(itY)]|}{|t|} dt + \sup_{\tau \in \mathbb{R}} \mathbb{P}[|Y - \tau| \le 1/T] \right).$$

Next we will require a multidimensional version of Esséen's concentration inequality.

Theorem 4.8 ([24, Lemma 7.17]) *There exists an absolute constant $C = C_{4.8} > 0$ such that the following statement holds. Given a random variable X in $\mathbb{R}^d$, we have that*

$$\sup_{\tau \in \mathbb{R}^d} \mathbb{P}[\|X - \tau\|_2 \le \varepsilon] \le \left(\frac{C_{4.8}\varepsilon}{\sqrt{d}} \right)^d \int_{\substack{\vec{\xi} \in \mathbb{R}^d \\ \|\vec{\xi}\|_2 \le d/\varepsilon}} |\mathbb{E}[\exp(2\pi i \vec{\xi} \cdot X)]| d\vec{\xi}.$$

Finally we will require the following consequence of Fourier inversion on lattices.

Theorem 4.9 *Given a bounded random variables $T \in \mathbb{Z}^d$ and $X \in \mathbb{R}$, possibly dependent, we have*

$$\mathbb{E}[\mathbb{1}_{T=\vec{t}} X] = (2\pi)^{-d} \int_{[-\pi,\pi]^d} \exp(-i\vec{t} \cdot \vec{\xi}) \mathbb{E}[X \exp(i\vec{\xi} \cdot T)] d\vec{\xi}.$$

5 Fourier coefficient bounds

5.1 Coefficient sequence

For the purposes of proving various central limit theorem and local central limit theorems, we will consider sums

$$\sum_{j=1}^{n} c_j \widetilde{a}_j$$

with coefficient sequences $(c_j)_{1 \le j \le n}$ which are more general than those arising from $\left(\sum_{1 \le k < j} \widetilde{b}_k + \frac{\widetilde{b}_j}{2} - (j - 1/2) \right)_{1 \le j \le n}$, which comes out of Lemma 2.3. The following definitions for such sequences arises from the proof; roughly, a sequence is *well-bounded* if it does not deviate much more than a simple random walk would, and it is *coarse* if it further resembles such a simple random walk at some finer scales.

Definition 5.1 We say a coefficient sequence $(c_j)_{1 \le j \le n}$ is *well-bounded* if the following conditions hold:

S1 $|c_j| \leq \sqrt{n} \log n$;

S2 $\sum_{j=1}^{n} c_j = 0$;

S3 $|c_j - c_k| \leq \sqrt{|j-k|} (\log n)^2$ for all $1 \leq j, k \leq n$;

and we say it is *coarse* if it is well-bounded and additionally the following hold:

S4 $\min_{a,b \in \mathbb{R}} \sum_{j=1}^{n} (c_j - aj - b)^2 \geq n^2/(\log n)^2$;

S5 There are at least $n/\log n$ indices j such that $c_j = c_{j+1} = c_{j+2} - 1/2$;

S6 For each integer $y \in [n^{1/4}, n/(\log n)^2]$ there are at least $n/\log n$ indices $1 \leq j \leq n - 2y$ such that $|c_j - 2c_{j+y} + c_{j+2y}| \geq \sqrt{y}$.

5.2 Fourier estimates

We now bound various Fourier expressions that will serve as a key input to our argument. We first define the basic setup.

Definition 5.2 Let Δ be a distribution which is either $\mathrm{Geom}(1/2)$ or $\mathrm{Pois}(1)$. Sample $X_j \sim \Delta$ independently for $1 \leq j \leq n$ and fix a sequence $(c_j)_{1 \leq j \leq n}$. Define the random variables

$$T_1 = \sum_{j=1}^{n} X_j - n, \qquad T_2 = \sum_{j=1}^{n} j X_j - \frac{n(n+1)}{2}, \qquad T_3 = 2 \sum_{j=1}^{n} c_j X_j.$$

We will be interested in Fourier coefficients of the form $\mathbb{E} \exp(i\vec{\Theta} \cdot (T_1, T_2, T_3))$. Our approach in general will be to reduce to essentially expressions involving Rademacher random variables and then to apply various basic bounds to conclude. (Note that we are not conditioning on the sum variable T_1 or "area" variable T_2 at this stage.)

Fact 5.3 *Given* $R \sim \mathrm{Ber}(1/2)$, $R \sim \mathrm{Geom}(1/2)$, *or* $R \sim \mathrm{Pois}(1)$ *and real* $|\Theta| \leq 3\pi/2$ *we have*

$$|\mathbb{E} \exp(iR\Theta)| \leq \exp(-c_{5.3}\Theta^2)$$

for some appropriate absolute constant $c_{5.3} > 0$.

Proof This follows immediately from the explicit computation that

$$|\mathbb{E} \exp(i\Theta R)| = \begin{cases} |\cos(\Theta/2)|, & \text{if } R \sim \mathrm{Ber}(1/2), \\ |2 - \exp(i\Theta)|^{-1}, & \text{if } R \sim \mathrm{Geom}(1/2), \\ \exp(\cos(\Theta) - 1), & \text{if } R \sim \mathrm{Pois}(1) \end{cases}$$

and some simple bounds based on Taylor series. $\square$

We first handle $\vec{\Theta}$ where $|\Theta_3|$ is large, since it is the most involved and serves as a basis for the other proofs. The key idea, which will be used to handle all the estimates present, is to extract independent random variables which isolate the effect of exactly one of the Θ_j.

Lemma 5.4 *Suppose that* $\vec{\Theta} = (\Theta_1, \Theta_2, \Theta_3)$ *is such that* $n^{-1/2}(\log n)^2 \leq |\Theta_3| \leq \pi$. *Then given Definition* 5.2 *and that* $(c_j)_{1 \leq j \leq n}$ *is coarse, we have*

$$|\mathbb{E}\exp(i\vec{\Theta} \cdot (T_1, T_2, T_3))| \leq n^{-\omega(1)}.$$

Remark This estimate, as well as Lemma 5.5, is only needed to establish Theorem 1.8.

Proof Since the coefficient sequence $(c_j)_{1 \leq j \leq n}$ is coarse, using **S5** there exists a 4-separated set of indices J (i.e., the difference of distinct elements is at least 4) such that $|J| = \Omega(n/\log n)$ and such that for $j \in J$ we have $c_j = c_{j+1} = c_{j+2} - 1/2$. We now claim that

$$(X_1, X_2, X_3) \overset{d.}{=} (1 - W)Z + W((1 - R)(0, 2, 0) + R(1, 0, 1)), \qquad (5.1)$$

where R, W, and Z are independent random variables defined via $R = \mathrm{Ber}(1/2)$,

$$W = \mathrm{Ber}\big(\min\{\mathbb{P}[(X_1, X_2, X_3) = (0, 2, 0)], \mathbb{P}[(X_1, X_2, X_3) = (1, 0, 1)]\}\big),$$

$$\mathbb{P}[Z = (k_1, k_2, k_3)] = \frac{1}{\mathbb{E}[1 - W]}$$

$$\cdot \left(\mathbb{P}[(X_1, X_2, X_3) = (k_1, k_2, k_3)] - \frac{1}{2}\mathbb{1}_{(k_1, k_2, k_3) \in \{(0,2,0),(1,0,1)\}}\mathbb{E}W\right)$$

for $(k_1, k_2, k_3) \in \mathbb{Z}_{\geq 0}^3$. Indeed, to see this let $2q = \min\{\mathbb{P}[(X_1, X_2, X_3) = (0, 2, 0)], \mathbb{P}[(X_1, X_2, \bar{X}_3) = (1, 0, 1)]\}$ and consider the following procedure: sample (X_1, X_2, X_3), but if either of the tuples $(x_1, x_2, x_3) \in \{(0, 2, 0), (1, 0, 1)\}$ is drawn then with probability $q/\mathbb{P}[(X_1, X_2, X_3) = (x_1, x_2, x_3)]$ enter a "resampling phase" where we decide with probability $1/2$ whether to output $(0, 2, 0)$ or $(1, 0, 1)$, overwriting the old value to produce a tuple (X_1', X_2', X_3'). (So, the "resampling phase" occurs with chance $2q$ by the law of total probability.) We see the distributional equality $(X_1', X_2', X_3') \overset{d.}{=} (X_1, X_2, X_3)$ by construction, but (X_1', X_2', X_3') is easily seen to be captured by the formula (5.1).

Note (5.1) holds even if we shift indices. Furthermore as the indices in J are 4-separated we have the distributional equality

$$((X_j, X_{j+1}, X_{j+2}))_{j \in J} \overset{d.}{=} ((1 - W_j)Z_j + W_j((0, 2, 0) + R_j(1, -2, 1)))_{j \in J}.$$

Notice by the triangle inequality and independence that

$$|\mathbb{E}\exp(i\vec{\Theta} \cdot (T_1, T_2, T_3))|$$

$$\leq \left|\mathbb{E}\left[\exp\left(i\sum_{j \in J+\{0,1,2\}}(\Theta_1 + \Theta_2 j + 2\Theta_3 c_j)X_j\right)\right]\right|$$

$$\leq \mathbb{E}_W\left[\left|\mathbb{E}\left[\exp\left(i\sum_{j \in J+\{0,1,2\}}(\Theta_1 + \Theta_2 j + 2\Theta_3 c_j)X_j\right)\Big|(W_j)_{j \in J}\right]\right|\right]$$

$$\leq \mathbb{E}_W\left[\left|\mathbb{E}\left[\exp\left(-i\sum_{j\in J}\mathbb{1}_{W_j=1}\Theta_3 R_j\right)\Big|(W_j)_{j\in J}\right]\right|\right]$$

$$\leq \mathbb{E}\left[\exp\left(-\Omega(\Theta_3^2)\sum_{j\in J}\mathbb{1}_{W_j=1}\right)\right]$$

$$\leq n^{-\omega(1)}.$$

The first and second lines follow from independence and the triangle inequality, the third follows from

$$(1,1,1)\cdot(1,-2,1)=0, \quad (j,j+1,j+2)\cdot(1,-2,1))=0,$$
$$(c_j,c_{j+1},c_{j+2})\cdot(1,-2,1)=-1/2,$$

and the fourth follows from Fact 5.3. In the final line we have used independence and Bernstein's inequality, which implies that $\#\{j\in J: W_j=1\}=\sum_{j\in J}\mathbb{1}_{W_j=1}\leq cn/\log n$ occurs with super-polynomially small probability for some small absolute constant $c>0$. $\qquad\square$

We next handle the case of intermediate $|\Theta_3|$. The remaining unhandled range will be in some sense controllable by an appropriate central limit theorem.

Lemma 5.5 *Suppose that* $\vec{\Theta}=(\Theta_1,\Theta_2,\Theta_3)$ *is such that* $n^{-1}(\log n)^3\leq|\Theta_3|\leq n^{-1/2}(\log n)^2$. *Then given Definition 5.2 and that* $(c_j)_{1\leq j\leq n}$ *is coarse, we have*

$$|\mathbb{E}\exp(i\vec{\Theta}\cdot(T_1,T_2,T_3))|\leq n^{-\omega(1)}.$$

Proof Let $y\in[n^{1/4},n/(\log n)^2]$ be an integer to be chosen later based on $n,|\Theta_3|$. Since $(c_j)_{1\leq j\leq n}$ is a coarse sequence, by **S6** there exists a set of indices J of size $\Omega(n/\log n)$ such that the sets $J,J+y,J+2y$ are disjoint and for each $j\in J$ we have $\sqrt{y}\leq|c_j-2c_{j+y}+c_{j+2y}|\leq 2\sqrt{y}(\log n)^2$ (the second inequality follows from two applications of **S3**). Therefore proceeding in an essentially identical manner to Lemma 5.4 (in particular writing $(X_j,X_{j+y},X_{j+2y})=(1-W_j)Z_j+W_j((0,2,0)+R_j(1,-2,1))$ for $j\in J$ similar to the proof of the previous lemma), we have that

$$|\mathbb{E}[\exp(i\vec{\Theta}\cdot(T_1,T_2,T_3))]|$$
$$\leq\left|\mathbb{E}\left[\exp\left(i\sum_{j\in J+\{0,y,2y\}}(\Theta_1+\Theta_2 j+2\Theta_3 c_j)X_j\right)\right]\right|$$
$$\leq\mathbb{E}\left[\left|\mathbb{E}\left[\exp\left(i\sum_{j\in J+\{0,1,2\}}(\Theta_1+\Theta_2 j+2\Theta_3 c_j)X_j\right)\Big|(W_j)_{j\in J}\right]\right|\right]$$
$$\leq\mathbb{E}\left[\left|\mathbb{E}\left[\exp\left(i\sum_{j\in J}\mathbb{1}_{W_j=1}\Theta_3 R_j(c_j-2c_{j+y}+c_{j+2y})\right)\Big|(W_j)_{j\in J}\right]\right|\right]$$

$$\leq \mathbb{E}\left[\exp\left(-\Omega(\Theta_3^2 y)\sum_{j\in J}\mathbb{1}_{W_j=1}\right)\right]$$

$$\leq n^{-\omega(1)}.$$

The reasoning is essentially identical to that in the proof of Lemma 5.4. We need that $|\Theta_3(c_j - 2c_{j+y} + c_{j+2y})| \leq \sqrt{y}(\log n)^2|\Theta_3| \leq 3\pi/2$ in order to apply Fact 5.3. If we additionally have that $\Theta_3^2 y \cdot n/(\log n) \geq (\log n)^2$, then using $|J| = \Omega(n/\log n)$ we can conclude the final estimate in a similar manner to the proof of Lemma 5.4. Thus it suffices to choose an integer y satisfying

$$(\log n)^3/(n\Theta_3^2) \leq y \leq 1/((\log n)^4\Theta_3^2)$$

and $y \in [n^{1/4}, n/(\log n)^2]$. This clearly exists by the given bounds for $|\Theta_3|$. □

We now prove a similar estimate for the case where $|\Theta_2|$ is near the maximum size. The proof is once again rather similar, but in this case we only need to consider consecutive pairs of indices $(j, j+1)$ in order to extract the necessary effect.

Lemma 5.6 *Suppose that* $\vec{\Theta} = (\Theta_1, \Theta_2, \Theta_3)$ *is such that* $n^{-1/2}\log n \leq |\Theta_2| \leq \pi$, $|\Theta_3| \leq n^{-1}(\log n)^3$ *and* $(c_j)_{1\leq j\leq n}$ *satisfies S1 and S3. Then given Definition 5.2, we have*

$$|\mathbb{E}\exp(i\vec{\Theta}\cdot(T_1, T_2, T_3))| \leq n^{-\omega(1)}.$$

Remark Lemmas 5.6–5.8 are needed for both Theorems 1.4 and 1.8. Note that these lemmas do not need an assumption on coarseness of $(c_j)_{1\leq j\leq n}$.

Proof Since the coefficient sequence $(c_j)_{1\leq j\leq n}$ satisfies **S3** we have $|c_j - c_{j+1}| \leq (\log n)^2$. Let $J \subseteq [n]$ be a 2-separated set of indices of size $\Omega(n)$. Furthermore note that $(X_1, X_2) \overset{d.}{=} (1-W)Z + W((1,0) + R(-1,1))$, where R, W, Z are independent random variables with $R = \mathrm{Ber}(1/2)$, $W = \mathrm{Ber}(\mathbb{P}[(X_1, X_2) = (1,0)])$, and

$$\mathbb{P}[Z = (k_1, k_2)]$$
$$= \frac{1}{\mathbb{E}[1-W]}\cdot\left(\mathbb{P}[(X_1, X_2, X_3) = (k_1, k_2)] - \frac{1}{2}\mathbb{1}_{(k_1,k_2)\in\{(0,1),(1,0)\}}\mathbb{E}W\right)$$

for $(k_1, k_2) \in \mathbb{Z}_{\geq 0}^2$, similar to as in the proof of Lemma 5.4.

Now for each index in J, we write $(X_j, X_{j+1}) = (1 - W_j)Z_j + W_j((1, 0) + R_j(-1, 1))$. Notice by the triangle inequality and independence that

$$
|\mathbb{E} \exp(i \vec{\Theta} \cdot (T_1, T_2, T_3))|
$$

$$
\leq \left| \mathbb{E} \left[\exp \left(i \sum_{j \in J + \{0, 1\}} (\Theta_1 + \Theta_2 j + 2\Theta_3 c_j) X_j \right) \right] \right|
$$

$$
\leq \mathbb{E} \left[\left| \mathbb{E} \left[\exp \left(i \sum_{j \in J + \{0, 1\}} (\Theta_1 + \Theta_2 j + 2\Theta_3 c_j) X_j \right) \Big| (W_j)_{j \in J} \right] \right| \right]
$$

$$
\leq \mathbb{E} \left[\left| \mathbb{E} \left[\exp \left(i \sum_{j \in J} \mathbb{1}_{W_j = 1} (\Theta_2 + 2(c_{j+1} - c_j)\Theta_3) R_j \right) \Big| (W_j)_{j \in J} \right] \right| \right]
$$

$$
\leq \mathbb{E} \left[\exp \left(-\Omega(\Theta_2^2) \sum_{j \in J} \mathbb{1}_{W_j = 1} \right) \right]
$$

$$
\leq n^{-\omega(1)}.
$$

The first and second line follows from triangle inequality, the third follows from $(1, 1) \cdot (-1, 1) = 0$, $(j, j+1) \cdot (-1, 1) = 1$, and $(c_j, c_{j+1}) \cdot (-1, 1) = c_{j+1} - c_j$ for $j \in J$, and the fourth from Fact 5.3 as well as $2|c_{j+1} - c_j||\Theta_3| \leq 2n^{-1}(\log n)^5 \leq |\Theta_2|/2$. In the final line we have once again used Bernstein's inequality. □

We next handle intermediate $|\Theta_2|$. In the remaining range central limit theorem type estimates become effective.

Lemma 5.7 *Suppose that $\vec{\Theta} = (\Theta_1, \Theta_2, \Theta_3)$ is such that $n^{-3/2}(\log n)^6 \leq |\Theta_2| \leq n^{-1/2} \log n$, $|\Theta_3| \leq n^{-1}(\log n)^3$ and $(c_j)_{1 \leq j \leq n}$ satisfies **S1** and **S3**. Then given Definition 5.2, we have*

$$
|\mathbb{E} \exp(i \vec{\Theta} \cdot (T_1, T_2, T_3))| \leq n^{-\omega(1)}.
$$

Proof Let $1 \leq y \leq n/8$ be an integer to be chosen later based on $n, |\Theta_2|$. Consider $J = \{\lfloor n/2 \rfloor - 2y, \lfloor n/2 \rfloor - 2y + 1, \ldots, \lfloor n/2 \rfloor - y\}$. We have

$$
|\Theta_3||c_j - c_{n-j}| \lesssim \sqrt{y}(\log n)^5/n, \quad 2y|\Theta_2| \leq |\Theta_2||n - 2j| \leq 4y|\Theta_2|
$$

for all $j \in J$. We ensure that y is chosen so that $\sqrt{y}(\log n)^5/n \leq cy|\Theta_2|$ for an appropriately small absolute constant $c > 0$ and so that $y|\Theta_2| \leq 1/8$ as well. We also guarantee $y \geq (\log n)^2$.

We now write $(X_j, X_{n-j}) = (1 - W_j)Z_j + W_j((1, 0) + R_j(-1, 1))$ for $j \in J$ in a similar manner to the proof of Lemma 5.6, and find

$$|\mathbb{E}\exp(i\vec{\Theta} \cdot (T_1, T_2, T_3))| \leq \left|\mathbb{E}\left[\exp\left(i\sum_{j \in J \cup (n-J)} (\Theta_1 + \Theta_2 j + 2\Theta_3 c_j) X_j\right)\right]\right|$$

$$\leq \mathbb{E}\left[\left|\mathbb{E}\left[\exp\left(i\sum_{j \in J \cup (n-J)} (\Theta_1 + \Theta_2 j + 2\Theta_3 c_j) X_j\right)\Big|(W_j)_{j \in J}\right]\right|\right]$$

$$\leq \mathbb{E}\left[\left|\mathbb{E}\left[\exp\left(i\sum_{j \in J} \mathbb{1}_{W_j=1}((n - 2j)\Theta_2 + 2(c_{n-j} - c_j)\Theta_3)R_j\right)\Big|(W_j)_{j \in J}\right]\right|\right]$$

$$\leq \mathbb{E}\left[\exp\left(-\Omega(y^2\Theta_2^2)\sum_{j \in J} \mathbb{1}_{W_j=1}\right)\right]$$

$$\leq n^{-\omega(1)}.$$

We used that $(n - 2j)\Theta_2$ dominates $2(c_{n-j} - c_j)\Theta_3$ in the second last line, as well as $2|(n - 2j)\Theta_2| \leq 1$ to apply Fact 5.3. For the last line, we note that $\sum_{j \in J} \mathbb{1}_{W_j=1} \lesssim y$ occurs with super-polynomially small probability (since $y \geq (\log n)^2$) and we are using the estimate $y^3\Theta_2^2 \geq (\log n)^2$.

To finish the proof, we check that it is possible to choose integer $1 \leq y \leq n/8$ with $y \geq (\log n)^{2/3}|\Theta_2|^{-2/3}$ and $y \geq (\log n)^2$ as well as $y \leq |\Theta_2|^{-1}/8$ and $y \geq c^{-2}(\log n)^{10}/(n^2\Theta_2^2)$. The bounds on $|\Theta_2|$ easily imply this is possible. $\square$

We now are finally in position to handle the cases where $|\Theta_1|$ is large. The remaining region will be handled by central limit theorem style techniques.

Lemma 5.8 *Suppose that $\vec{\Theta} = (\Theta_1, \Theta_2, \Theta_3)$ is such that $n^{-1/2}(\log n)^7 \leq |\Theta_1| \leq 5\pi/4$, $|\Theta_2| \leq n^{-3/2}(\log n)^6$, and $|\Theta_3| \leq n^{-1}(\log n)^3$ and $(c_j)_{1 \leq j \leq n}$ satisfies S1 and S3. Then given Definition 5.2, we have*

$$|\mathbb{E}\exp(i\vec{\Theta} \cdot (T_1, T_2, T_3))| \leq n^{-\omega(1)}.$$

Proof Note that $|\Theta_2|n + |\Theta_3| \max_{1 \leq j \leq n} |c_j| \leq 2(\log n)^6 n^{-1/2}$. Therefore we have

$$|\mathbb{E}[\exp(\vec{\Theta} \cdot (T_1, T_2, T_3))]| = \left|\mathbb{E}\left[\exp\left(i\sum_{j=1}^n (\Theta_1 + \Theta_2 j + 2\Theta_3 c_j) X_j\right)\right]\right|$$

$$\leq \exp(-\Omega(n\Theta_1^2)) \leq n^{-\omega(1)},$$

where we have simply noted that Θ_1 dominates $\Theta_2 j + 2\Theta_3 c_j$ and applied Fact 5.3. $\square$

We now prove the desired estimate for the region which is approximately within the region which is controlled via a central limit theorem. For completeness we provide

a short proof via an argument closely related to the Lindeberg exchange method [16] and the proof of the Gaussian invariance principle [19] (see Theorem 4.5). This will help us reduce computing the necessary integrals to a purely Gaussian integration problem.

Lemma 5.9 *Suppose that* $\vec{\Theta} = (\Theta_1, \Theta_2, \Theta_3)$ *is such that* $|\Theta_1| \leq n^{-1/2}(\log n)^7$, $|\Theta_2| \leq n^{-3/2}(\log n)^6$, $|\Theta_3| \leq n^{-1}(\log n)^3$, *and* $(c_j)_{1 \leq j \leq n}$ *satisfies **S1** and **S2**. Given Definition 5.2, we further define*

$$T_1' = \sum_{j=1}^n X_j', \qquad T_2' = \sum_{j=1}^n j X_j', \qquad T_3' = 2 \sum_{j=1}^n c_j X_j',$$

where we independently sample $X_j' \sim \mathcal{N}(0, \mathrm{Var}[\Delta])$. *Then we have*

$$|\mathbb{E}[\exp(i\vec{\Theta} \cdot (T_1, T_2, T_3))] - \mathbb{E}[\exp(i\vec{\Theta} \cdot (T_1', T_2', T_3'))]| \lesssim n^{-1/2}(\log n)^{21}.$$

Remark *If* $\Delta = \mathrm{Geom}(1/2)$ *then* $(\mathbb{E}[\Delta], \mathrm{Var}[\Delta]) = (1, 2)$, *and if* $\Delta = \mathrm{Pois}(1)$ *then* $(\mathbb{E}[\Delta], \mathrm{Var}[\Delta]) = (1, 1)$.

Proof Notice that by iteratively replacing X_i by X_i' and applying the triangle inequality we have

$$|\mathbb{E} \exp(i\vec{\Theta} \cdot (T_1, T_2, T_3)) - \mathbb{E} \exp(i\vec{\Theta} \cdot (T_1', T_2', T_3'))|$$

$$\leq \sum_{j=1}^n \left| \left(\mathbb{E} \exp(i\vec{\Theta} \cdot (1, j, 2c_j)(X_j - 1)) - \mathbb{E} \exp\left(i\vec{\Theta} \cdot (1, j, 2c_j)X_j'\right) \right) \right.$$

$$\left. \times \mathbb{E} \exp\left(i\vec{\Theta} \cdot \left(\sum_{1 \leq j' < j} (1, j', 2c_{j'})X_{j'}' + \sum_{j < j' \leq n} (1, j', 2c_{j'})(X_{j'} - 1) \right) \right) \right|$$

$$\leq \sum_{j=1}^n |\mathbb{E} \exp(i\vec{\Theta} \cdot (1, j, 2c_j)(X_j - 1)) - \mathbb{E} \exp(i\vec{\Theta} \cdot (1, j, 2c_j)X_j')|$$

$$\lesssim \sum_{j=1}^n (|\Theta_1| + n|\Theta_2| + 2n^{1/2}(\log n)|\Theta_3|)^3 \mathbb{E}[|X_j|^3 + |X_j'|^3] \lesssim n^{-1/2}(\log n)^{21}.$$

To justify the second-to-last inequality, we use that $|\exp(ix) - 1 - ix + x^2/2| \leq |x|^3$ for $x \in \mathbb{R}$ from Taylor's theorem and that the first and second moments of $X_j - 1$ and X_j' match. $\qquad \square$

6 Translating Fourier information

We now translate Fourier information into probabilistic information in order to prove Theorem 1.4. We defer the proof of the following lemma, which shows that certain

coefficient sequences that will arise in our computation are well-bounded with very high probability, until the next section.

Lemma 6.1 *Fix m and let $1 \leq k^* \leq m$. Consider $\Theta \neq 0$ and $\theta = (\theta_{jk})_{1 \leq j < k \leq m}$ with $\|\theta\|_\infty \leq \Theta$. Consider independent random variables $X_j^{(k)} \sim \Delta$ for $1 \leq k \leq m$ and $1 \leq j \leq n$, where $\Delta \in \{\mathrm{Geom}(1/2), \mathrm{Pois}(1)\}$. Finally, let*

$$c_j^{(k^*)} = \frac{1}{2\Theta} \left(\sum_{k<k^*} \theta_{kk^*} \left(\sum_{j'=1}^n (M_n^*)_{jj'} \left(X_{j'}^{(k)} - 1 \right) \right) \right.$$
$$\left. + \sum_{k>k^*} \theta_{k^*k} \left(\sum_{j'=1}^n (M_n^*)_{j'j} \left(X_{j'}^{(k)} - 1 \right) \right) \right).$$

Then with probability $1 - n^{-\omega(1)}$ we have that $(c_j^{(k^*)})_{1 \leq j \leq n}$ is well-bounded (Definition 5.1).

Now we prove Theorem 1.4, which we recall for convenience.

Theorem 1.4 *Fix $m \geq 2$ and independently sample n-sided dice $A_1, \ldots, A_m$, either all from the multiset model or all from the balanced sequence model. Let $G^{(j)}$ for $1 \leq j \leq m$ be infinite vectors of standard Gaussians and for $1 \leq j < k \leq m$ let*

$$H_{jk} = \sum_{\ell \geq 1} \sigma_\ell (G_{2\ell-1}^{(j)} G_{2\ell}^{(k)} - G_{2\ell}^{(j)} G_{2\ell-1}^{(k)}).$$

Then for any digraph D on vertices $[m]$,

$$\lim_{n \to \infty} \mathbb{P}[A_j \text{ beats } A_k \text{ for all } jk \in E(D)] = \mathbb{P}[H_{jk} > 0 \text{ for all } jk \in E(D)].$$

Proof of Theorem 1.4 given Lemma 6.1 Sample k dice either all from the multiset model or all from the balanced sequence model, $A_1, \ldots, A_m$ with $A_j = (a_{j1}, \ldots, a_{jn}) \in [n]^n$. Let the frequency counts of die A_k be $\widetilde{a}_{ki} = |\{j : a_{kj} = i\}|$. We are given the tournament D on $[k]$ and wish to understand the chance that A_i beats A_j precisely when ij is a directed edge. (Note that we may assume D is a full tournament since partial tournaments clearly follow by summing appropriately.)

If we are in the multiset model let $\Delta = \mathrm{Geom}(1/2)$ and if we are in the balanced sequence model let $\Delta = \mathrm{Pois}(1)$. Consider m independent copies of the setup in Definition 5.2 (ignoring the sequence c_j and random variable T_3), denoted by $(X_j^{(k)})_{1 \leq j \leq n}$ and $(T_j^{(k)})_{1 \leq j \leq 2}$ for $1 \leq k \leq m$, corresponding to the k-th die. Note that $(\widetilde{a}_{ki})_{1 \leq i \leq n}$ is distributed as $(X_j^{(k)})_{1 \leq j \leq n}$ conditional on $T_1^{(k)} = T_2^{(k)} = 0$ by Lemma 2.2.

Let

$$Y_{k_1 k_2} := X^{(k_2)\mathsf{T}} M_n^* X^{(k_1)} = (X^{(k_2)} - \sqrt{n}\vec{v}_1)^\mathsf{T} M_n^* (X^{(k_1)} - \sqrt{n}\vec{v}_1)$$

for $1 \leq k_1 < k_2 \leq m$ (recall $\vec{v}_1$ from Definition 2.4), and for $\theta = (\theta_{k_1 k_2})_{1 \leq k_1 < k_2 \leq m}$ let

$$Y(\theta) := \exp\left(i \sum_{1 \leq k_1 < k_2 \leq m} \theta_{k_1 k_2} Y_{k_1 k_2}\right).$$

By Theorem 4.9 with $T := (T_b^{(k)})_{k \in [m], b \in [2]}$ and indexing the coordinates of $\vec{\xi}$ by $(\xi_{kb})_{k \in [m], b \in [2]}$, we have

$$\mathbb{E}[\mathbb{1}_{T=\vec{0}} Y(\theta)] = (2\pi)^{-2m} \int_{[-\pi,\pi]^{2m}} \mathbb{E}[Y(\theta) \exp(i\vec{\xi} \cdot T)] d\vec{\xi}$$

$$= (2\pi)^{-2m} \int_{[-\pi,\pi]^{2m}} \mathbb{E}\left[\exp\left(i \sum_{1 \leq k_1 < k_2 \leq m} \theta_{k_1 k_2} Y_{k_1 k_2} + i\vec{\xi} \cdot T\right)\right] d\vec{\xi}.$$

$$(6.1)$$

We fix some θ satisfying $\|\theta\|_\infty \leq n^{-1}(\log n)^3$. Given this condition, we will now estimate the integrand and show that it is very small unless $\|\vec{\xi}_{\cdot 1}\|_\infty = \tilde{O}(n^{-1/2})$ and $\|\vec{\xi}_{\cdot 2}\|_\infty = \tilde{O}(n^{-3/2})$.

We now collect terms so as to express the argument in the exponential as a linear function of $X^{(k^*)}$ with coefficients depending on $X^{(k)}$ for $k \neq k^*$. We see

$$\sum_{1 \leq k_1 < k_2 \leq m} \theta_{k_1 k_2} Y_{k_1 k_2} + \vec{\xi} \cdot T$$

$$= \sum_{j=1}^n \left(\sum_{k < k^*} \theta_{kk^*} \left(\sum_{j'=1}^n (M_n^*)_{jj'} \left(X_{j'}^{(k)} - 1\right)\right)\right.$$

$$+ \sum_{k > k^*} \theta_{k^*k} \left(\sum_{j'=1}^n (M_n^*)_{j'j} \left(X_{j'}^{(k)} - 1\right)\right)\right) X_j^{(k^*)} + \tilde{Y}_{k^*}$$

$$+ \xi_{k^*,1} \sum_{j=1}^n \left(X_j^{(k^*)} - 1\right) + \xi_{k^*,2} \sum_{j=1}^n j \left(X_j^{(k^*)} - 1\right) \qquad (6.2)$$

for some $\tilde{Y}_{k^*}$ that depends only on $(X^{(k)})_{k \neq k^*}$. Consider $\Theta = (\log n)^3/n$, and define

$$c_j^{(k^*)} := \frac{1}{2\Theta} \left(\sum_{k < k^*} \theta_{kk^*} \left(\sum_{j'=1}^n (M_n^*)_{jj'} \left(X_{j'}^{(k)} - 1\right)\right) + \sum_{k > k^*} \theta_{k^*k} \left(\sum_{j'=1}^n (M_n^*)_{j'j} \left(X_{j'}^{(k)} - 1\right)\right)\right)$$

for $k^* \in [m]$ and $j \in [n]$. We have

$$
\left| \mathbb{E} \exp\left(i \sum_{1 \le k_1 < k_2 \le m} \theta_{k_1 k_2} Y_{k_1 k_2} + i\vec{\xi} \cdot T \right) \right|
$$

$$
= \left| \mathbb{E} \exp\left(i \left(\Theta \sum_{j=1}^{n} 2c_j^{(k^*)} X_j^{(k^*)} + \xi_{k^*,1} \sum_{j=1}^{n} (X_j^{(k^*)} - 1) \right. \right. \right.
$$

$$
\left. \left. \left. + \xi_{k^*,2} \sum_{j=1}^{n} j(X_j^{(k^*)} - 1) + \widetilde{Y}_{k^*} \right) \right) \right|. \tag{6.3}
$$

Now we apply Lemmas 5.6–5.8 to gain control over ξ. In order to use these, we need each $(c_j^{(k^*)})_{1 \le j \le n}$ for $k^* \in [m]$ to be a well-bounded coefficient sequence. By Lemma 6.1, this occurs with probability $1 - n^{-\omega(1)}$ over $(X^{(k)})_{k \ne k^*}$.

So if $n^{-1/2} \log n \le |\xi_{k^*2}| \le \pi$ then by Lemma 5.6 we have that (6.3) is of magnitude $n^{-\omega(1)}$: condition on an outcome of $(X^{(k)})_{k \ne k^*}$ for which $c_j^{(k^*)}$ is well-bounded using Lemma 6.1, and then apply Lemma 5.6. We are using that $\Theta = n^{-1}(\log n)^3$. Similarly, if $n^{-3/2}(\log n)^6 \le |\xi_{k^*2}| \le n^{-1/2} \log n$ then by Lemmas 5.7 and 6.1 we see that (6.3) is of magnitude $n^{-\omega(1)}$. Finally, if $|\xi_{k^*2}| \le n^{-3/2}(\log n)^6$ and $n^{-1/2}(\log n)^7 \le |\xi_{k^*1}| \le \pi$ then Lemmas 5.8 and 6.1 show (6.3) is of magnitude $n^{-\omega(1)}$.

Combining these observations with (6.1) and (6.2) we see

$$
\mathbb{E}[\mathbb{1}_{T=\vec{0}} Y(\theta)] = (2\pi)^{-2m} \int_{[-\tau_1, \tau_1]^m \times [-\tau_2, \tau_2]^m}
$$

$$
\mathbb{E}\left[\exp\left(i \sum_{1 \le k_1 < k_2 \le m} \theta_{k_1 k_2} Y_{k_1 k_2} + i\vec{\xi} \cdot T \right) \right] d\vec{\xi} \pm n^{-\omega(1)}, \tag{6.4}
$$

where $\tau_1 = n^{-1/2}(\log n)^7$ and $\tau_2 = n^{-3/2}(\log n)^6$. Additionally, the product in the region of integration is interpreted as corresponding to the choice of $b \in \{1, 2\}$, i.e., the region is defined by $|\xi_{k1}| \le \tau_1$ and $|\xi_{k2}| \le \tau_2$.

Recall also that we assumed $\|\theta\|_\infty \le n^{-1}(\log n)^3$. We can now use an approach similar to the proof of Lemma 5.9 (or [16, 19]) to exchange the variables $X_j^{(k)}$ with shifted Gaussians $Z_j^{(k)} + 1$, where $Z_j^{(k)} \sim \mathcal{N}(0, \text{Var}[\Delta])$. Note that

$$
\sum_{1 \le k_1 < k_2 \le m} \theta_{k_1 k_2} Y_{k_1 k_2} + \vec{\xi} \cdot T = \sum_{1 \le k_1 < k_2 \le m} \theta_{k_1 k_2} \left(X^{(k_2)} - \sqrt{n}\vec{v}_1 \right)^\mathsf{T} M_n^* (X^{(k_1)} - \sqrt{n}\vec{v}_1)
$$

$$
+ \sum_{k=1}^{m} \xi_{k1} \sum_{j=1}^{n} \left(X_j^{(k)} - 1 \right) + \sum_{k=1}^{m} \xi_{k2} \sum_{j=1}^{n} j(X_j^{(k)} - 1). \tag{6.5}
$$

Now since $X_j^{(k)} - 1$ are independent and mean 0, and have variance $\mathrm{Var}[\Delta]$, we are in position to apply Theorem 4.5. We first compute that the influences for the degree 2 multilinear polynomial corresponding to (6.5) are bounded by

$$O((\|M_n^*\|_{1\to2}^2 + \|M_n^{*\intercal}\|_{1\to2}^2)\cdot\|\theta\|_\infty^2 + \|\xi_{\cdot1}\|_\infty^2 + n^2\|\xi_{\cdot2}\|_\infty^2) = O(n^{-1}(\log n)^{14})$$

using Lemma 2.6 (specifically, **M4**).

Let $Z_j^{(k)} \sim \mathcal{N}(0, \mathrm{Var}[\Delta])$ be independent Gaussians and let

$$\widetilde{Z}(\vec{\xi}) := \sum_{1\le k_1<k_2\le m} \theta_{k_1 k_2} Z^{(k_2)\intercal} M_n^* Z^{(k_1)} + \sum_{k=1}^m \xi_{k1}\sum_{j=1}^n Z_j^{(k)} + \sum_{k=1}^m \xi_{k2}\sum_{j=1}^n j Z_j^{(k)}.$$

By Theorem 4.5 and (6.4) we have

$$\mathbb{E}[\mathbb{1}_{T=\vec{0}}Y(\theta)] = (2\pi)^{-2m}\int_{[-\tau_1,\tau_1]^m\times[-\tau_2,\tau_2]^m}\mathbb{E}[\exp(i\widetilde{Z}(\vec{\xi}))]d\vec{\xi}$$
$$\pm O((\tau_1\tau_2)^m n^{-1/2}(\log n)^{21}). \tag{6.6}$$

Note that the latter two sums in $\widetilde{Z}$, which involve $\vec{\xi}$, only depend on $Z^{(k)}\cdot\vec{v}_1$ and $Z^{(k)}\cdot\vec{v}_2$ whereas the bilinear forms only depend on the projection of $Z^{(k)}$ to the orthogonal complement of $\mathrm{span}_\mathbb{R}\{\vec{v}_1, \vec{v}_2\}$ (by Definition 2.4). Therefore we see that the first sum is independent from the latter two. This means that the integrand in (6.6) is the product of some constant and some multivariate Gaussian characteristic function.

Now, if $\|\xi_{\cdot1}\|_\infty \ge \tau_1$ or $\|\xi_{\cdot2}\|_\infty \ge \tau_2$, then easily we find there is some $k^* \in [m]$ with

$$\sum_{j=1}^n (\xi_{k^*1} + j\xi_{k^*2})^2 \gtrsim (\log n)^{12}.$$

We therefore deduce that for such $\vec{\xi}$,

$$|\mathbb{E}[\exp(i\widetilde{Z}(\vec{\xi}))]| = \left|\mathbb{E}\exp\left(i\sum_{1\le k_1<k_2\le m}\theta_{k_1 k_2} Z^{(k_2)\intercal} M_n^* Z^{(k_1)}\right)\right|\cdot\exp(-\Omega((\log n)^{12}))$$
$$\le n^{-\omega(1)}.$$

Furthermore, since the integrand is proportional to the characteristic function of some multivariate Gaussian, it is easy to see that the integral to infinity over such $\vec{\xi}$ is still $n^{-\omega(1)}$ in size. So, from (6.6) we deduce

$$\mathbb{E}[\mathbb{1}_{T=\vec{0}}Y(\theta)] = (2\pi)^{-2m}\int_{\mathbb{R}^{2m}}\mathbb{E}[\exp(i\widetilde{Z}(\vec{\xi}))]d\vec{\xi} \pm O((\tau_1\tau_2)^m n^{-1/2}(\log n)^{21})$$

$$= \mathbb{E} \exp \left(i \sum_{1 \le k_1 < k_2 \le m} \theta_{k_1 k_2} Z^{(k_2)\mathsf{T}} M_n^* Z^{(k_1)} \right)$$

$$\times (2\pi)^{-2m} \int_{\mathbb{R}^{2m}} \mathbb{E} \exp \left(i \left(\sum_{k=1}^{m} \xi_{k1} \sum_{j=1}^{n} Z_j^{(k)} + \sum_{k=1}^{m} \xi_{k2} \sum_{j=1}^{n} j Z_j^{(k)} \right) \right) d\vec{\xi}$$

$$\pm O((\tau_1 \tau_2)^m n^{-1/2} (\log n)^{21}).$$

Plugging in $\theta = \vec{0}$ and dividing, and noting that the integral in the last line is order $\Theta((n^{-1/2} \cdot n^{-3/2})^m)$ (treating m as fixed), we deduce

$$\mathbb{E}[Y(\theta)|T = \vec{0}] = \mathbb{E} \exp \left(i \sum_{1 \le k_1 < k_2 \le m} \theta_{k_1 k_2} Z^{(k_2)\mathsf{T}} M_n^* Z^{(k_1)} \right)$$

$$\pm O(n^{-1/2} (\log n)^{21+13m}) \tag{6.7}$$

for $\|\theta\|_\infty \le n^{-1} (\log n)^3$.

We wish to show

$$\left(\frac{Y_{jk}}{n \mathrm{Var}[\Delta]} \middle| T = 0 \right)_{1 \le j < k \le m} \xrightarrow{d.} (H_{jk})_{1 \le j < k \le m}$$

since Lemma 2.5 (and the facts $\vec{v}_1^\mathsf{T} M_n^* = 0$ and $M_n^* \vec{v}_1 = 0$) shows A_j beats A_k precisely when $Y_{jk} > 0$. Now let $G^{(j)}$ and H_{jk} be as in Theorem 1.4. From (6.7) and Lévy continuity, we see it is enough to show

$$\left(\frac{Z^{(k)\mathsf{T}} M_n^* Z^{(j)}}{n \mathrm{Var}[\Delta]} \right)_{1 \le j < k \le m} \xrightarrow{d.} (H_{jk})_{1 \le j < k \le m}$$

as $n \to \infty$. (Simple inspection of the proof shows that this would also imply the second remark following Theorem 1.4.) Note that we may assume $\mathrm{Var}[\Delta] = 1$ since $Z_\ell^{(j)} \sim \mathcal{N}(0, \mathrm{Var}[\Delta])$ and we are now in a scale-invariant situation with respect to Δ.

We are now purely in a setting of joint convergence of certain bilinear forms of standard Gaussian vectors. Thus, the problem will ultimately reduce to limiting spectral properties of the operators M_n^*. By a variant of the spectral theorem, since M_n^* is skew-symmetric by Lemma 2.6 (**M2**), we can write $M_n^* = Q_n \Sigma_n Q_n^\mathsf{T}$, where Q_n is orthogonal and Σ_n consists of diagonal 2×2 blocks of the form

$$\begin{bmatrix} 0 & -\sigma_{n,\ell} \\ \sigma_{n,\ell} & 0 \end{bmatrix}$$

for $\ell \in [\lfloor n/2 \rfloor]$, and possibly a single 0 in the final diagonal entry if n is odd. By orthogonal invariance of Gaussian vectors, applying the orthogonal matrix Q_n, our

distribution is the same as

$$\left(\frac{G^{(k)\mathsf{T}}\Sigma_n G^{(j)}}{n}\right)_{1\le j<k\le m},$$

where $G^{(j)}$ are independent standard Gaussian vectors. We have

$$\frac{G^{(k)\mathsf{T}}\Sigma_n G^{(j)}}{n} = \sum_{\ell=1}^{\lfloor n/2\rfloor}\frac{\sigma_{n,\ell}}{n}(G^{(j)}_{2\ell-1}G^{(k)}_{2\ell} - G^{(j)}_{2\ell}G^{(k)}_{2\ell-1}). \tag{6.8}$$

We have that for any constant $t \ge 1$, $(\sigma_{n,\ell}/n)_{1\le\ell\le t} \to (\sigma_\ell)_{1\le\ell\le t}$ as $n \to \infty$ by Lemma 2.6 (**M8**).

Now consider some fixed $t \ge 1$ (which we will take to be growing slowly at the end of this argument). Using $\sum_{\ell\ge t}(\sigma_{n,t}/n)^2 = O(1/t)$ and Chebyshev's inequality we easily see that with probability $1 - O(t^{-1/2})$, the sum in (6.8) over indices $\ell \ge t$ contributes at most $O(t^{-1/4})$. Furthermore, $(\sigma_{n,\ell}/n)_{1\le\ell\le t} \to (\sigma_\ell)_{1\le\ell\le t}$ as $n \to \infty$ by the above argument. Hence, we deduce that with probability at least $1 - O(t^{-1/2})$,

$$\left(\frac{G^{(k)\mathsf{T}}\Sigma_n G^{(j)}}{n}\right)_{1\le j<k\le m}$$

is within ℓ^∞ distance $O(t^{-1/4})$ of a random vector which converges to

$$\left(\sum_{\ell=1}^{t}\sigma_\ell\left(G^{(j)}_{2\ell-1}G^{(k)}_{2\ell} - G^{(j)}_{2\ell}G^{(k)}_{2\ell-1}\right)\right)_{1\le j<k\le m}$$

in distribution. Finally, taking $t \to \infty$ slowly gives the desired result, recalling from the remark following Theorem 1.4 that almost surely the appropriate sums converge as $t \to \infty$. $\qquad\square$

7 Properties of coefficient sequences

We next prove Lemma 6.1.

Proof of Lemma 6.1 By definition we have

$$\sum_{j=1}^{n}(M_n^*)_{jj'} = \sum_{j=1}^{n}(M_n^*)_{j'j} = 0$$

hence $\sum_{j=1}^{n}c_j^{(k^*)} = 0$ immediately follows, establishing **S2**. Note also that $c_j^{(k^*)}$ is a weighted sum of independent random variables $X-1$, where $X \sim \Delta$. Since $\mathbb{E}[\Delta] = 0$ and Δ is either Poisson or geometric we easily see that it is a sum of independent mean

0 random variables with bounded $\|X - 1\|_{\psi_1}$. Additionally, the coefficients of $c_j^{(k^*)}$ are of the form $\theta_{kk^*}(M_n^*)_{jj'}/(2\Theta)$ and $\theta_{k^*k}(M_n^*)_{j'j})/(2\Theta)$, which by definition and Lemma 2.6 (**M3**) are bounded in magnitude.

Hence we can apply Bernstein's inequality (Theorem 4.1) to obtain

$$\mathbb{P}[|c_j^{(k^*)}| \geq t] \leq 2\exp\left(-c_{4.1}\min\left(\frac{t^2}{O(n)}, \frac{t}{O(1)}\right)\right).$$

Choose $t = \sqrt{n}\log n$, which implies that the event $|c_j^{(k^*)}| \geq \sqrt{n}\log n$ occurs with probability at most $\exp(-\Omega((\log n)^2))$. Taking a union bound over n events for $1 \leq j \leq n$, we obtain **S1** with probability $1 - n^{-\omega(1)}$.

Finally, **S3** is similar. We wish to show $|c_{j_1} - c_{j_2}| \leq \sqrt{|j_1 - j_2|}(\log n)^2$ for all $1 \leq j_1 < j_2 \leq n$ occurs with probability $1 - n^{-\omega(1)}$, as then a union bound will finish the proof. To do this, we will exploit cancellations in $(M_n^*)_{j_1 j'} - (M_n^*)_{j_2 j'}$. In particular, we have

$$c_{j_1}^{(k^*)} - c_{j_2}^{(k^*)} = \frac{1}{2\Theta}\left(\sum_{k<k^*}\theta_{kk^*}\left(\sum_{j'=1}^{n}\left((M_n^*)_{j_1 j'} - (M_n^*)_{j_2 j'}\right)\left(X_{j'}^{(k)} - 1\right)\right)\right.$$
$$\left. + \sum_{k>k^*}\theta_{k^*k}\left(\sum_{j'=1}^{n}\left((M_n^*)_{j' j_1} - (M_n^*)_{j' j_2}\right)(X_{j'}^{(k)} - 1)\right)\right).$$

By Lemma 2.6 (**M7**) we have that $|(M_n^*)_{j_1 j'} - (M_n^*)_{j_2 j'}| = O(|j_1 - j_2|/n)$ for all but $O(|j_1 - j_2|)$ values of j', for which the value is $O(1)$. Since M_n^* is skew-symmetric (**M2**), the same occurs when we transpose the matrix. Therefore we can use Bernstein's inequality (Theorem 4.1) again, this time deducing

$$\mathbb{P}[|c_{j_1}^{(k^*)} - c_{j_2}^{(k^*)}| \geq t] \leq 2\exp\left(-c_{4.1}\min\left(\frac{t^2}{O(|j_1 - j_2|)}, \frac{t}{O(1)}\right)\right).$$

Taking $t = \sqrt{|j_1 - j_2|}(\log n)^2 \geq (\log n)^2$ and taking a union bound, we deduce the desired conclusion. □

We now prove that the coefficient sequence coming from Lemma 2.3 is typically coarse. This will be used to prove Theorem 1.8 later. We note that the idea of breaking into various intervals and extracting tuples of coefficients with the desired properties also appears in the work of Polymath, in particular in [22, Lemma 5.10]; however the proofs here are simpler as we require only a "physical space" condition on the coefficients.

Lemma 7.1 *Let* $\Delta \in \{\text{Geom}(1/2), \text{Pois}(1)\}$. *Let* $\widetilde{X}_j \sim \Delta$ *for all* $1 \leq j \leq n$ *and then condition on* $\sum_{j=1}^{n} \widetilde{X}_j = n$ *and* $\sum_{j=1}^{n} j\widetilde{X}_j = n(n+1)/2$. *If*

$$c_j = \sum_{1 \leq k < j}\widetilde{X}_j + \frac{\widetilde{X}_j}{2} - (j - 1/2)$$

then with probability $1 - n^{-\omega(1)}$ the sequence $(c_j)_{1 \le j \le n}$ is coarse (Definition 5.1).

Proof We will prove that everything but **S2** occurs in the unconditioned independent model with probability $1 - n^{-\omega(1)}$. Then note that

$$\mathbb{P}\left[\sum_{j=1}^{n} \widetilde{X}_j = n \wedge \sum_{j=1}^{n} j\widetilde{X}_j = \frac{n(n+1)}{2}\right] \gtrsim n^{-2}$$

from Lemma 9.1 (which is proved only using results up to Sect. 5) or from the line before (6.7) in the proof of Theorem 1.4.

Thus the failure probability of any property in the conditional model will be at most equal to $(n^{-\omega(1)})/(\Omega(n^{-2})) = n^{-\omega(1)}$ by Bayes' rule. So it suffices to consider the independent model, noting that **S2** follows from the conditions $\sum_{j=1}^{n} \widetilde{X}_j = n$ and $\sum_{j=1}^{n} j\widetilde{X}_j = n(n+1)/2$.

S1 and **S3** are simple Bernstein inequality calculations, similar to the proof of Lemma 6.1, and we omit the details. For **S5**, note that $c_j = c_{j+1} = c_{j+2} - 1/2$ follows if $\widetilde{X}_j = \widetilde{X}_{j+1} = 0$ and $\widetilde{X}_{j+2} = 1$. Let J be a 3-separated sequence of size $\Omega(n)$ and note that $j \in J$ satisfies the condition required by **S5** with probability $\Omega(1)$. Thus Bernstein's inequality or Chernoff easily implies **S5**.

For **S6**, consider J which is all multiples of $4y$ in $\{1, \ldots, n - 4y\}$, of size $\Omega(n/y)$. For each $j \in J$, the probability that $|c_j - 2c_{j+y} + c_{j+2y}| \ge \sqrt{y}$ is seen to be $\Omega(1)$ by the central limit theorem, and this is independent over all $j \in J$. Thus by Bernstein or Chernoff, with probability at least $1 - \exp(-\Omega(n/y))$ there are at least $\Omega(n/y)$ many $j \in J$ satisfying the condition required by **S6**. We can repeat the argument for the translations of J by $\{1, 2, \ldots, y\}$ and take a union bound, which yields $\Omega(n)$ many indices j with probability $1 - n^{-\omega(1)}$ as desired.

Finally, we consider **S4**. We can mimic the proof of **S6** above except with $y = \lfloor n/(\log n)^{3/2} \rfloor$ and still deduce that with probability $1 - n^{-\omega(1)}$, there are at least $\Omega(n)$ indices $1 \le j \le n - 2y$ with $|c_j - 2c_{j+y} + c_{j+2y}| \ge \sqrt{y}$. We can pass to a subset J of size $\Omega(n)$ with the property that $j - j' \notin \{\pm y, \pm 2y\}$ for all $j, j' \in J$. For each $j \in J$ we have

$$(c_j - aj - b)^2 + (c_{j+y} - a(j+y) - b)^2 + (c_{j+2y} - a(j+2y) - b)^2$$
$$\ge \frac{1}{4}(c_j - 2c_{j+y} + c_{j+2y})^2$$

using the inequality $x_1^2 + x_2^2 + x_3^2 \ge (x_1 - 2x_2 + x_3)^2/4$. Hence we deduce

$$\sum_{j=1}^{n}(c_j - aj - b)^2 \gtrsim |J| \cdot (\sqrt{y})^2/4 \gtrsim ny \ge n^2/(\log n)^2$$

for all $a, b \in \mathbb{R}$. The result follows. $\qquad\square$

8 Consequences of Theorem 1.4

We now derive the claimed symmetry facts from the statement of Theorem 1.4.

Proof of Corollary 1.5 For the first consequence, note that the Gaussian distribution is negation invariant and therefore the result for reversing the edges at vertex u follows by negating the Gaussian $G^{(u)}$ in Theorem 1.4. For the second consequence, simple replace every die with its "complement", i.e., we map $(a_1, \ldots, a_n)$ to $(n + 1 - a_n, \ldots, n + 1 - a_i)$. (In the limiting expression of Theorem 1.4, this corresponds to switching $G^{(j)}_{2\ell-1}$ and $G^{(j)}_{2\ell}$ for all $1 \leq j \leq m$ and $\ell \geq 1$.) □

Now we turn to Corollary 1.7. We require the following lemma relating a tournamenton having image in the set $\{0, 1\}$ to the distribution of its k-vertex subtournaments. Recall that for tournamenton $\mathcal{T}$, a $\mathcal{T}$-random tournament on M vertices is obtained by sampling M random points $x_1, \ldots, x_M$ from $[0, 1]$ uniformly at random, and then sampling a directed edge from x_i to x_j with probability $\mathcal{T}(x_i, x_j)$ (and otherwise putting one from x_j to x_i) for all $1 \leq i < j \leq M$.

Lemma 8.1 *Fix a tournamenton $\mathcal{T}$. Suppose that for every $\varepsilon > 0$, for all M sufficiently large there is a set $\mathcal{F}_M$ of M-vertex tournaments with $|\mathcal{F}_M| \leq 2^{\varepsilon M^2}$ such that a $\mathcal{T}$-random tournament on M vertices lies in $\mathcal{F}_M$ with probability at least $1 - \varepsilon$. Then $\mu(\{(x, y) \colon \mathcal{T}(x, y) \notin \{0, 1\}\}) = 0$, where μ is the Lebesgue measure on $[0, 1]^2$.*

Proof Suppose that $\mu(\{(x, y) \colon \mathcal{T}(x, y) \notin \{0, 1\}\}) > 0$. Then there exists $\delta > 0$ such that

$$\mu(\{(x, y) \colon \mathcal{T}(x, y) \in [\delta, 1 - \delta]\}) \geq \delta. \tag{8.1}$$

Sample an M-vertex $\mathcal{T}$-random tournament, and define $x_1, \ldots, x_M$ as above. Let

$$X_M = \{(i, j) \in [M]^2 \colon i < j \text{ and } \mathcal{T}(x_i, x_j) \in [\delta, 1 - \delta]\}.$$

We have that $\mathbb{E}|X_M| \geq \delta\binom{M}{2}$ from (8.1) and thus by applying the Azuma–Hoeffding inequality (Lemma 4.2) on the Doob martingale formed by revealing $x_1, \ldots, x_M$ in order, we see that $\mathbb{P}[X_M \geq \delta M^2/4] \geq 1 - \delta$ for M sufficiently large as a function of δ.

This means there is an event $\mathcal{E}$ occurring with probability at least $1 - \delta$ over the randomness of $x_1, \ldots, x_M$ such that conditional on $\mathcal{E}$, the entropy of our $\mathcal{T}$-random tournament is at least $H(\text{Ber}(\delta)) \cdot |X_M| \gtrsim \delta^2 M^2$.

But by initial assumption there is an event $\mathcal{F}$ holding with probability $1 - \varepsilon$ such that the original M-vertex tournament is in $\mathcal{F}_M$. We see that the entropy of the $\mathcal{T}$-random tournament must be at most $H(\varepsilon) + \log_2 |\mathcal{F}_M| + \varepsilon \log_2(2^{M^2}) \lesssim \varepsilon M^2$. Taking ε much smaller than δ^2 and M sufficiently large, we obtain a contradiction. □

We now are in position to prove Corollary 1.7.

Proof of Corollary 1.7 We first note that T_n converges to a limit tournamenton $\mathcal{T}$ since Theorem 1.4 implies that for a fixed digraph D the associated densities converge. Thus the result follows via convergence of subgraph densities implying convergence in cut metric (see [9] where this theory is worked out in the case of directed graphs; the theory for tournamentons follows as a direct consequence via say applying [25, Theorem 4.1] which characterizes a directed graph limit being a tournamenton in terms of certain subgraph counts vanishing).

The more difficult part of Corollary 1.7 is verifying the conditions of Lemma 8.1. Fix m dice, where we will consider m large, and consider the random series H_{jk} from Theorem 1.4. For these m dice, reveal $G_\ell^{(j)}$ for $\ell \le 2\lfloor m^{1/2} \rfloor$ and round the value to the nearest $1/m^{25}$, and label each vertex with the corresponding tuple of values. Call the collection of these labels $L(G)$, which depends only on $G_\ell^{(j)}$ for $\ell \le 2\lfloor m^{1/2} \rfloor$. Note that with probability $1 - \exp(-\Omega(m))$ all these sampled Gaussians are bounded by m and hence there is a set $\mathcal{L}$ of at most $\exp(O(m^{3/2} \log m))$ different possible labelings such that $L(G) \in \mathcal{L}$ under this event. Furthermore given these labels $L(G)$, the value

$$H_{jk}^* = \sum_{\ell=1}^{\lfloor m^{1/2} \rfloor} \sigma_\ell \left(G_{2\ell-1}^{(j)} G_{2\ell}^{(k)} - G_{2\ell}^{(j)} G_{2\ell-1}^{(k)} \right)$$

is pinned down to within an interval $I_{jk}(G)$ (defined whenever $L(G) \in \mathcal{L}$) of length at most m^{-20}, say, for all $1 \le j < k \le m$.

Note that $H_{j,k} - H_{j,k}^*$ has variance $O(m^{-1/2})$ and hence with probability $1 - \exp(-m^{\Omega(1)})$ all these infinite tails are of magnitude at most say $m^{-1/5}$ by Theorem 4.6.

Let $\mathcal{L}'$ be the set of labelings $L(G) \in \mathcal{L}$ such that the interval $I_{jk}(G)$ intersects $[-m^{-1/5}, m^{1/5}]$ for at most $m^{2-1/20}$ many choices of $1 \le j < k \le m$. Let $B(G)$ be the set of (j, k) where there is an intersection. Note $B(G)$ depends only on $L(G)$ whenever $L(G) \in \mathcal{L}$. Combining the observations above, there is an event $\mathcal{E}$ which occurs with probability $1 - \exp(-m^{\Omega(1)})$ such that the following holds if we assume $\mathcal{E}$:

- $L(G) \in \mathcal{L}$, where $\mathcal{L}$ is a deterministic set of size $\exp(O(m^{3/2} \log m))$;
- If $L(G) \in \mathcal{L}'$ then the digraph $D(G) := \{(j, k) : H_{jk} > 0\}$ depends only on the identity of $L(G)$ and on whether (j, k) or (k, j) is in $D(G)$ for all $(j, k) \in B(G)$. Here $\mathcal{L}'$ is the deterministic subset of $\mathcal{L}$ defined above.

If we can show that $L(G) \in \mathcal{L}'$ with probability $1 - O(m^{-1/20})$, say, then by Theorem 1.4 this will establish the hypotheses of Lemma 8.1 and hence this will finish the proof. Indeed, then we know that with high probability the digraph $D(G)$ can be determined by revealing $L(G) \in \mathcal{L}$ (with $\exp(O(m^{3/2} \log m))$ choices), which determines $B(G)$, and then revealing whether $(j, k) \in D(G)$ for all $(j, k) \in B(G)$, which has at most $2^{m^{2-1/20}}$ choices. This will establish the hypothesis of Lemma 8.1 for $M = \Omega(\varepsilon^{-20})$, say.

Finally, by Theorem 4.3 for fixed $1 \le j < k \le m$ the probability that $H_{jk}^* = O(m^{-1/5})$ is at most $O(m^{-1/10})$. Therefore by Markov's inequality, there are at most $m^{2-1/20}$ pairs in $B(G)$ with probability $1 - O(m^{-1/20})$. The result follows. $\qquad\square$

We now give a sketch of an alternative proof of Corollary 1.7 pointed out by the anonymous referee; this proof more directly stems from Theorem 1.4.

Sketch of alternative proof of Corollary 1.7 By using measure-isomorphisms note that one can view digraphons as measurable functions $W\colon X^2 \to [0, 1]$ where X is $\Omega = (X, \mu)$ is any atomless probability space. Consider $X = \mathbb{R}^{\mathbb{N}^+}$ equipped with the Borel σ-algebra (of the product topology). Let μ be the distribution of an infinite vector of standard Gaussians and define $W\colon X^2 \to [0, 1]$ by

$$W(x, y) = \mathbb{1}\left[\sum_{\ell \geq 1} \sigma_\ell(x_{2\ell-1}y_{2\ell} - x_{2\ell}y_{2\ell-1}) > 0\right].$$

By Theorem 1.4, sampling a digraph on m vertices from W has desired densities. Therefore $\mathcal{T}$ is exactly W by the fact that convergence of tournamentons in cut metric is equivalent to convergence of digraph homomorphism densities. Furthermore W is obviously a $\{0, 1\}$-tournamenton and thus the desired result follows. $\square$

We end by providing a short proof given Corollary 1.7 that the four-cycle (A beats B, B beats C, C beats D, and D beats A) occurs with a greater than $1/16$ limiting probability. This consequence was the method used by Cornacchia and Hązła to disprove quasirandomness of dice tournaments (in the simpler model where die faces are drawn independently at random from the uniform distribution on $[0, 1]$). The surprising fact that this limiting probability is larger than $1/16$ falls out naturally of Corollarys 1.5 and 1.7. Note that this means that if A beats B, B beats C, and C beats D then D is *more* likely to beat A in the limit, since a path with 3 edges is a tree so has limiting probability $1/8$. (An analogous result holds for larger even cycles which we leave as an exercise to the reader.)

Proposition 8.2 *Let $\mathcal{T}$ be as in Corollary 1.7. We have that*

$$\int_0^1 \int_0^1 \int_0^1 \int_0^1 \mathcal{T}(x_1, x_2)\mathcal{T}(x_2, x_3)\mathcal{T}(x_3, x_4)\mathcal{T}(x_4, x_1)dx_1dx_2dx_3dx_4 > \frac{1}{16}.$$

Proof Note that

$$\int_0^1 \int_0^1 \int_0^1 \int_0^1 \mathcal{T}(x_1, x_2)\mathcal{T}(x_2, x_3)\mathcal{T}(x_3, x_4)\mathcal{T}(x_4, x_1)dx_1dx_2dx_3dx_4$$

$$= \int_0^1 \int_0^1 \int_0^1 \int_0^1 \mathcal{T}(x_1, x_2)\mathcal{T}(x_2, x_3)\mathcal{T}(x_4, x_3)\mathcal{T}(x_1, x_4)dx_1dx_2dx_3dx_4$$

$$= \int_0^1 \int_0^1 \left(\int_0^1 \mathcal{T}(x_1, x_2)\mathcal{T}(x_2, x_3)dx_2\right)^2 dx_1dx_3$$

$$\geq \left(\int_0^1 \int_0^1 \int_0^1 \mathcal{T}(x_1, x_2)\mathcal{T}(x_2, x_3)dx_1dx_3dx_2\right)^2$$

$$= \left(\int_0^1 \int_0^1 \int_0^1 \mathcal{T}(x_2, x_1)\mathcal{T}(x_2, x_3)dx_1dx_3dx_2\right)^2$$

$$= \left(\int_0^1 \left(\int_0^1 \mathcal{T}(x_2, x_1)dx_1 \right)^2 dx_2 \right)^2$$

$$\geq \left(\int_0^1 \int_0^1 \mathcal{T}(x_2, x_1)dx_1dx_2 \right)^4 = \frac{1}{16},$$

where we have applied Corollary 1.5 on vertex 4, factoring the square, Cauchy–Schwarz, Corollary 1.5 on vertex 1, factoring the square, Cauchy–Schwarz, and then used that $\mathcal{T}$ has average $1/2$ (from Theorem 1.4 and symmetry). For equality to occur we must have that for almost all x, y,

$$1/4 = \int_0^1 \mathcal{T}(x, z)\mathcal{T}(z, y)dz.$$

By the equivalence of codegree counts with quasirandomness for tournaments; see [6, P4, Theorem 1] in work of Chung and Graham, in order for equality to occur we must have $\mathcal{T}(x, y) = 1/2$ almost everywhere in Lebesgue measure. This contradicts the statement of Corollary 1.7, so the inequality is strict. □

We also briefly derive that any tournament T on m vertices occurs in the limit with positive probability. This recovers a recent result of Akin [1] (which was proven by dynamical methods and which in turn reproves results of Moon and Moser [18] which allows for dice to not have the same means and a result of Finkelstein and Thorp [10] which constructs cycles of arbitrary length via a more explicit construction).

Proposition 8.3 *Recall the setup of Theorem 1.4 and fix any digraph D. We have that*

$$\mathbb{P}[H_{jk} > 0 \text{ for all } jk \in E(D)] > 0.$$

Equivalently, the D-density in $\mathcal{T}$ is positive. In particular, given any D there exists a set of dice which produce the digraph D.

Proof Let C be a sufficiently large constant to be chosen later. Let $|E(D)| = u$ and label the edges of the digraph D by $e_1, \ldots, e_u$ and the vertices by $1, \ldots, m$. We may assume D is connected so $m \leq u + 1$. We correspond the indices $\{2\ell - 1, 2\ell\}$ to directed edge e_ℓ. For each $\ell \in [u]$ and each $i \in [m]$ which is not an endpoint of the edge e_ℓ we define the event $\mathcal{E}_{\ell,i}$:

$$\max \left\{ |G_{2\ell-1}^{(i)}|, |G_{2\ell}^{(i)}| \right\} \leq 1.$$

For each $\ell \in [u]$, if e_ℓ is an edge directed from j to k then we define the event $\mathcal{E}_\ell$:

$$G_{2\ell-1}^{(j)}, G_{2\ell}^{(k)} \geq Cu^{1/2}, \qquad |G_{2\ell}^{(j)}|, |G_{2\ell-1}^{(k)}| \leq 1.$$

Recall that

$$H_{jk} = \sum_{\ell \geq 1} \sigma_\ell \left(G_{2\ell-1}^{(j)} G_{2\ell}^{(k)} - G_{2\ell}^{(j)} G_{2\ell-1}^{(k)} \right).$$

Let us further define the event $\mathcal{E}_{\text{tail}}$:

$$\left| \sum_{\ell > |E(D)|} \sigma_\ell \left(G_{2\ell-1}^{(j)} G_{2\ell}^{(k)} - G_{2\ell}^{(j)} G_{2\ell-1}^{(k)} \right) \right| \leq C$$

for all $jk \in E(D)$. By Chebyshev's inequality, for any given $j, k \in [m]$ this occurs with probability $1 - O(1/(C^2 u))$. Taking a union bound over $jk \in E(D)$ we see $\mathbb{P}[\mathcal{E}_{\text{tail}}] \geq 1/2$ if C is large enough.

Note that $\mathcal{E}_{\ell,i}$, $\mathcal{E}_\ell$, $\mathcal{E}_{\text{tail}}$ are jointly independent, so they jointly occur with positive probability (at least $\exp(-O(u^2))$).

If jk is directed edge $e_{\ell*}$, then by construction the first $|E(D)|$ terms of H_{jk} contribute $\gtrsim C^2 \sigma_{\ell*} u$, so $\Omega(C^2)$ by Lemma 2.6 (**M9**). On the other hand, the tail $\ell > |E(D)|$ contributes at most C. Thus $H_{jk} > 0$ if C was chosen large enough. The result follows. □

9 Proof of Theorem 1.8

Finally we compute the probability of having a tie. We proceed in a slightly indirect manner via first considering the probability that a given *coarse die* (i.e., an appropriate associated sequence is coarse in the sense of Definition 5.1) ties with a randomly sampled die. This amounts to computing the chance that $T_1 = T_2 = 0$ and $T_1 = T_2 = T_3 = 0$ given the setup of Definition 5.2, which will be the first step in understanding the necessary probability.

Lemma 9.1 *Assume the setup of Definition 5.2 and that $(c_j)_{1 \leq j \leq n}$ is coarse. We have*

$$\mathbb{P}[T_1 = 0 \wedge T_2 = 0] = \frac{\sqrt{3}}{\pi \operatorname{Var}[\Delta] n^2} + O(n^{-5/2} (\log n)^{34}),$$

$$\mathbb{P}[T_1 = 0 \wedge T_2 = 0 \wedge T_3 = 0]$$

$$= \frac{\sqrt{3}}{(2\pi \operatorname{Var}[\Delta])^{3/2} n^2 (\min_{a,b \in \mathbb{R}} \sum_{j=1}^n (c_j - aj - b)^2)^{1/2}} + O(n^{-7/2} (\log n)^{37}).$$

Proof Let

$$\widetilde{T_1} = \sum_{j=1}^n \widetilde{X}_j, \quad \widetilde{T_2} = \sum_{j=1}^n j \widetilde{X}_j, \quad \widetilde{T_3} = 2 \sum_{j=1}^n c_j \widetilde{X}_j,$$

where $\widetilde{X}_j \sim \mathcal{N}(0, \operatorname{Var}[\Delta])$ are independent Gaussians. Let $\vec{X} \in \mathbb{R}^n$ be the vector with these coordinates. Define the sets

$$R_2 = \{(\xi_1, \xi_2) \colon |\xi_1| \leq n^{-1/2} (\log n)^7, \ |\xi_2| \leq n^{-3/2} (\log n)^6\},$$

$$R_3 = \{(\xi_1, \xi_2, \xi_3) \colon |\xi_1| \leq n^{-1/2} (\log n)^7, \ |\xi_2| \leq n^{-3/2} (\log n)^6, \ |\xi_3| \leq n^{-1} (\log n)^3\}.$$

By Theorem 4.9 and Lemmas 5.4 to 5.8 we easily see

$$\mathbb{P}[T_1 = 0 \wedge T_2 = 0] = (2\pi)^{-2} \int_{R_2} \mathbb{E}[\exp(i\vec{\xi} \cdot (T_1, T_2))]d\vec{\xi} \pm n^{-\omega(1)},$$

$$\mathbb{P}[T_1 = 0 \wedge T_2 = 0 \wedge T_3 = 0] = (2\pi)^{-3} \int_{R_3} \mathbb{E}[\exp(i\vec{\xi} \cdot (T_1, T_2, T_3))]d\vec{\xi} \pm n^{-\omega(1)}.$$

Then, using Lemma 5.9 to transfer to Gaussians we find

$$\mathbb{P}[T_1 = 0 \wedge T_2 = 0] = (2\pi)^{-2} \int_{R_2} \mathbb{E}[\exp(i\vec{\xi} \cdot (\widetilde{T}_1, \widetilde{T}_2))]d\vec{\xi} + O(n^{-5/2}(\log n)^{34}),$$

$$(9.1)$$

$$\mathbb{P}[T_1 = 0 \wedge T_2 = 0 \wedge T_3 = 0] = (2\pi)^{-3} \int_{R_3} \mathbb{E}[\exp(i\vec{\xi} \cdot (\widetilde{T}_1, \widetilde{T}_2, \widetilde{T}_3))]d\vec{\xi}$$
$$+ O(n^{-7/2}(\log n)^{37}).$$

$$(9.2)$$

We define the matrix $3 \times n$ matrix Q_3 via

$$Q_3 := \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 2 & \dots & n \\ 2c_1 & 2c_2 & \dots & 2c_n \end{pmatrix}$$

and we let Q_2 be the first two rows of Q_3. We have $\vec{\xi} \cdot (\widetilde{T}_1, \widetilde{T}_2, \widetilde{T}_3) = \vec{\xi}^\mathsf{T} Q_3 \vec{X}$ and $\mathbb{E}[\vec{\xi} \cdot (\widetilde{T}_1, \widetilde{T}_2, \widetilde{T}_3)] = 0$. Furthermore note that $\mathbb{E}[(\vec{\xi}^\mathsf{T} Q_3 \vec{X})^2] = \mathrm{Var}[\Delta] \cdot (\vec{\xi}^\mathsf{T} Q_3 Q_3^\mathsf{T} \vec{\xi})$. We deduce

$$\mathbb{P}[T_1 = 0 \wedge T_2 = 0 \wedge T_3 = 0] = (2\pi)^{-3} \int_{R_3} \exp(-\mathrm{Var}[\Delta] \cdot \xi^\mathsf{T} Q_3 Q_3^\mathsf{T} \vec{\xi}/2))d\vec{\xi}$$
$$+ O(n^{-7/2}(\log n)^{37}),$$

$$\mathbb{P}[T_1 = 0 \wedge T_2 = 0] = (2\pi)^{-2} \int_{R_2} \exp(-\mathrm{Var}[\Delta] \cdot \vec{\xi}^\mathsf{T} Q_2 Q_2^\mathsf{T} \vec{\xi}/2)d\vec{\xi}$$
$$+ O(n^{-5/2}(\log n)^{34}).$$

We compute

$$Q_3 Q_3^\mathsf{T} = \begin{pmatrix} n & n(n+1)/2 & 0 \\ n(n+1)/2 & n(n+1)(2n+1)/6 & 2\sum_{j=1}^{n} jc_j \\ 0 & 2\sum_{j=1}^{n} jc_j & 4\sum_{j=1}^{n} c_j^2 \end{pmatrix},$$

recalling **S2** which implies $(1, \ldots, 1) \cdot (c_1, \ldots, c_n) = 0$. Also, since $Q_3^\mathsf{T} \vec{e}_1$, $Q_3^\mathsf{T} \vec{e}_3$ are orthogonal (denoting $\vec{e}_j \in \mathbb{R}^3$ as the jth elementary vector), we deduce

$$
\begin{aligned}
\mathrm{dist}(Q_3^\mathsf{T} \vec{e}_2, \ \mathrm{span}_\mathbb{R}\{Q_3^\mathsf{T} \vec{e}_j\}_{j \in \{1,3\}})^2 \\
&= \mathrm{dist}((-(n-1)/2, -(n-3)/2, \ldots, (n-1)/2), \mathrm{span}_\mathbb{R}\{(c_1, \ldots, c_n)\})^2 \\
&\gtrsim n^3 \left(1 - \frac{\langle (-(n-1)/2, -(n-3)/2, \ldots, (n-1)/2), (c_1, \ldots, c_n)\rangle^2}{\|(-(n-1)/2, -(n-3)/2, \ldots, (n-1)/2)\|_2^2 \|(c_1, \ldots, c_n)\|_2^2} \right) \\
&\gtrsim n^3 \cdot \frac{\min_{a,b \in \mathbb{R}} \sum_{j=1}^n (c_j - aj - b)^2}{\sum_{j=1}^n c_j^2} \\
&\gtrsim n^3/(\log n)^4.
\end{aligned}
$$

The second-to-last line comes from noting that the desired minimum corresponds to the distance from $(c_1, \ldots, c_n)$ to the plane spanned by $Q_3^\mathsf{T} \vec{e}_1$, $Q_3^\mathsf{T} \vec{e}_2$. The last line uses **S1** and **S4**. Similarly, we find

$$
\mathrm{dist}(Q_3^\mathsf{T} \vec{e}_3, \ \mathrm{span}_\mathbb{R}\{Q_3^\mathsf{T} \vec{e}_j\}_{j \in \{1,2\}})^2 = \min_{a,b \in \mathbb{R}} \sum_{j=1}^n (c_j - aj - b)^2 \geq n^2/(\log n)^2.
$$

Therefore we have

$$
\begin{aligned}
\xi^\mathsf{T} Q_3 Q_3^\mathsf{T} \xi = \|Q_3^\mathsf{T} \xi\|_2^2 &\geq \max_{j^* \in [3]} \xi_{j^*}^2 \mathrm{dist}(M^\mathsf{T} \vec{e}_{j^*}, \mathrm{span}_\mathbb{R}(\{M\vec{e}_j\}_{j \in [3] \setminus \{j^*\}})^2 \\
&\gtrsim n\xi_1^2 + n^3 \xi_2^2/(\log n)^4 + n^2 \xi_3^2/(\log n)^2.
\end{aligned}
$$

This inequality immediately allows us to extend the regions of integration in (9.1) and (9.2) to $\mathbb{R}^2$ and $\mathbb{R}^3$, respectively, since the integrand within the remaining region is super-polynomially small and decaying rapidly. Applying the formula for a Gaussian integral, we have

$$
\begin{aligned}
\mathbb{P}[T_1 = 0 \wedge T_2 = 0] &= (2\pi)^{-2} \int_{\mathbb{R}^2} \exp(-\mathrm{Var}[\Delta] \cdot \vec{\xi}^\mathsf{T} Q_2 Q_2^\mathsf{T} \vec{\xi}/2) d\vec{\xi} + O(n^{-5/2}(\log n)^{34}) \\
&= (2\pi)^{-1} (\det Q_2 Q_2^\mathsf{T})^{-1/2} + O(n^{-5/2}(\log n)^{34}) \\
&= \frac{\sqrt{3}}{\pi \mathrm{Var}[\Delta] n^2} + O(n^{-5/2}(\log n)^{34}).
\end{aligned}
$$

We computed the determinant explicitly as $n^2(n^2 - 1)/12$ using the expression for $Q_3 Q_3^\mathsf{T}$. Similarly, we have

$$\mathbb{P}[T_1 = 0 \wedge T_2 = 0 \wedge T_3 = 0] = (2\pi)^{-3} \int_{\mathbb{R}^3} \exp(-\text{Var}[\Delta] \cdot \vec{\xi}^\mathsf{T} Q_3 Q_3^\mathsf{T} \vec{\xi}/2) d\vec{x}$$

$$+ O(n^{-7/2}(\log n)^{37})$$

$$= (2\pi \text{Var}[\Delta])^{-3/2} \det(Q_3 Q_3^\mathsf{T})^{-1/2} + O(n^{-7/2}(\log n)^{37})$$

$$= \frac{\sqrt{3}}{(2\pi \text{Var}[\Delta])^{3/2} n^2 (\min_{a,b \in \mathbb{R}} \sum_{j=1}^n (c_j - aj - b)^2)^{1/2}} + O(n^{-7/2}(\log n)^{37}).$$

In the last line we used the base times height formula row-by-row to compute $(\det Q_3 Q_3^\mathsf{T})^{1/2}$, which can be interpreted as the 3-dimensional volume of the corresponding parallelepiped spanned by $Q_3^\mathsf{T} \vec{e}_1, Q_3^\mathsf{T} \vec{e}_2, Q_3^\mathsf{T} \vec{e}_3$ within $\mathbb{R}^n$. $\qquad\square$

Given Lemma 9.1, and recalling Lemma 2.2, the approach will now be to take an average of $\mathbb{P}[T_3 = 0 | T_1 = T_2 = 0]$ over the distribution of coarse sequences $(c_j)_{1 \leq j \leq n}$ that come from the frequency count statistics of a typical die sampled from either model. This requires us to understand the quadratic expressions $\min_{a,b \in \mathbb{R}} \sum_{j=1}^n (c_j - aj - b)^2$ (where c_j will linearly depend on the frequency count statistics) conditional on stuff such as $T_1 = T_2 = 0$. At a high level, we will reduce understanding a quadratic form to understanding finitely many linear forms jointly (via sampling random rows to take a dot product against; heuristically, one could study the large singular vectors). Thus we will reduce to a situation where we only need the sort of Fourier control guaranteed by Sect. 5. Specifically, we can prove Lemma 9.4 given the tools in Sect. 5, which is the key estimate. Beyond this, we need various tools to control certain tail estimates and related notions that occur in the course of the proof, which is very much related to the fact that evaluating $\mathbb{E}(\sum_{\ell \geq 1} \sigma_\ell^2 (Z_\ell^2 + Z_\ell'^2))^{-1/2}$ involves a (convergent) improper integral.

We will require the following estimate regarding sampling independent points for a given distribution on $[n]$. We use the following estimate on sums of independent random variables from [27].

Theorem 9.2 ([27, Theorem 4]) *Fix $\beta \in [1, 2]$. There exists $C_{9.2}(\beta) > 0$ such that the following holds. Let X_i be independent mean zero random variables with $\mathbb{E}|X_i|^\beta < \infty$. We have that*

$$\mathbb{E}\left[\left|\sum_{i=1}^n X_i\right|^\beta\right] \leq C_{9.2}(\beta) \sum_{i=1}^n \mathbb{E}|X_i|^\beta$$

Lemma 9.3 *There exists a constant $C_{9.3} > 0$ such that the following holds. Given a sequence $x_1, \ldots, x_n$, let $i_1, \ldots, i_M$ be indices chosen uniformly at random from $[n]$. Then*

$$\left|\frac{1}{M}\sum_{j=1}^M x_{i_j} - \frac{1}{n}\sum_{j=1}^n x_j\right| \leq C_{9.3} M^{-1/4} \left(\frac{\sum_{j=1}^n |x_j|^{3/2}}{n}\right)^{2/3}$$

occurs with probability at least $1 - M^{-1/8}$.

Proof We have that

$$
\mathbb{E}\left|\frac{1}{M}\sum_{j=1}^{M}x_{i_j}-\frac{1}{n}\sum_{j=1}^{n}x_j\right|^{3/2}=M^{-3/2}\mathbb{E}\left|\sum_{j=1}^{M}\left(x_{i_j}-\frac{1}{n}\sum_{j=1}^{n}x_j\right)\right|^{3/2}
$$

$$
\lesssim M^{-1/2}\cdot\mathbb{E}\left|x_{i_1}-\frac{1}{n}\sum_{j=1}^{n}x_j\right|^{3/2}
$$

$$
\lesssim M^{-1/2}\cdot\mathbb{E}|x_{i_1}-x_{i_2}|^{3/2}\lesssim M^{-1/2}\cdot\mathbb{E}|x_{i_1}|^{3/2}
$$

$$
=n^{-1}M^{-1/2}\sum_{i=1}^{n}|x_i|^{3/2},
$$

where we have used Theorem 9.2, Jensen's inequality, and that $|x-y|^{3/2}\leq 2^{1/2}(|x|^{3/2}+|y|^{3/2})$. The desired follows immediately by Markov's inequality. □

We next verify the following key distributional identity. We will use it after applying the identity $\min_{a,b\in\mathbb{R}}\sum_j(c_j-a-bj)^2=\|M_n^*\widetilde{b}\|_2^2$, where $c_j=\sum_{1\leq k<j}\widetilde{b}_k+\frac{b_j}{2}-(j-1/2)$. This fact comes from the definition of M_n^* in Definition 2.4 as the result of applying two projection operations to M_n in the orthogonal directions $\vec{v}_1,\vec{v}_2\in\mathbb{R}^n$ which span the same space as $(1,1,\ldots,1),(1,2,\ldots,n)$; see (9.9).

Lemma 9.4 *Let X_j, T_1, T_2 be as in Definition 5.2, sample $\widetilde{X}_j\sim\mathcal{N}(0,\mathrm{Var}[\Delta])$ for $1\leq j\leq n$, let $X=(X_1,\ldots,X_n)$, and let $\widetilde{X}=(\widetilde{X}_1,\ldots,\widetilde{X}_n)$. If $M\leq\log\log n$ and $j_k\in[n]$ and $|\Theta_k|\leq(\log n)^{7/4}$ for $1\leq k\leq M$ then*

$$
\left|\mathbb{E}\left[\exp\left(i\sum_{k=1}^{M}\frac{\Theta_k\langle X,M_n^*\vec{e}_{j_k}\rangle}{\sqrt{n}}\right)\bigg|T_1=T_2=0\right]\right.
$$

$$
\left.-\mathbb{E}\left[\exp\left(i\sum_{k=1}^{M}\frac{\Theta_k\langle\widetilde{X},M_n^*\vec{e}_{j_k}\rangle}{\sqrt{n}}\right)\right]\right|\leq n^{-1/2}(\log n)^{39}.
$$

Proof The proof is very similar to the first part of the proof of Theorem 1.4 as well as the proof of Lemma 9.1, so we will be brief and focus only on the necessary modifications from the basic proof strategy.

We apply Theorem 4.9 to deduce

$$
\mathbb{E}\left[\mathbb{1}_{T_1=T_2=0}\exp\left(i\sum_{k=1}^{M}\frac{\Theta_k\langle X,M_n^*\vec{e}_{j_k}\rangle}{\sqrt{n}}\right)\right]
$$

$$
=\frac{1}{(2\pi)^2}\int_{[-\pi,\pi]^2}\mathbb{E}\exp\left(i\sum_{k=1}^{M}\frac{\Theta_k\langle X,M_n^*\vec{e}_{j_k}\rangle}{\sqrt{n}}+i\vec{\xi}\cdot(T_1,T_2)\right)d\vec{\xi}
$$

and then apply Lemmas 5.6–5.8. In order to apply these, we define

$$c_j^* := n \sum_{k=1}^{M} \frac{\Theta_k (M_n^* \vec{e}_{j_k})_j}{2\sqrt{n}}$$

so that

$$\sum_{k=1}^{M} \frac{\Theta_k \langle X, M_n^* \vec{e}_{j_k} \rangle}{\sqrt{n}} = \frac{2 \sum_{j=1}^{n} c_j^* X_j}{n}.$$

However, $(c_j^*)_{1 \leq j \leq n}$ does not quite satisfy **S3** so we cannot apply the lemmas directly; the only obstruction is that c_j^* will have slight "local jumps" near $j \in \{j_1, \ldots, j_M\}$ due to the fact that M_n^* has slightly "discontinuous" entries along the diagonal. Indeed, one can check

$$|c_j^* - c_{j'}^*| \lesssim \frac{|j - j'|}{\sqrt{n}} (\log n)^{7/4} \log \log n + \sqrt{n} \sum_{k=1}^{M} \mathbb{1}_{j_k \in [j, j']}$$

for all $1 \leq j < j' \leq n$ due to **M7**. This is at most $\sqrt{|j - j'|}(\log n)^2$ if $\{j_1, \ldots, j_M\} \cap [j, j'] = \emptyset$.

To fix this obstruction, we simply apply Lemmas 5.6–5.8 to a consecutive sequence of entries. Note that there is some $\{n_1 + 1, n_1 + 2, \ldots, n_2\} \subseteq [n]$ with $|n_2 - n_1| \geq n/(2M)$ with no j_k contained in this consecutive range. Let $n' = |n_2 - n_1|$. Therefore, we may condition on values of $X_{[n] \setminus \{n_1+1, \ldots, n_2\}}$ and then deduce the necessary estimate from the randomness of $X_{\{n_1+1, \ldots, n_2\}}$. We deduce

$$\left| \mathbb{E} \exp \left(i \sum_{k=1}^{M} \frac{\Theta_k \langle X, M_n^* \vec{e}_{j_k} \rangle}{\sqrt{n}} + i\vec{\xi} \cdot (T_1, T_2) \right) \right| \leq n^{-\omega(1)}$$

as long as $|\xi_2| \in [(n')^{-3/2} (\log n')^6, \pi]$ or $|\xi_2| \leq (n')^{-3/2} (\log n')^6$ and $|\xi_1 + n_1 \xi_2| \in [(n')^{-1/2} (\log n')^7, 5\pi/4]$ (where one applies Lemmas 5.6–5.8). This trivially covers all (ξ_1, ξ_2) except for say $|\xi_1| \leq n^{-1/2} (\log n)^8$ and $|\xi_2| \leq n^{-3/2} (\log n)^7$.

Then we can apply Lemma 5.9 (technically, since ξ_1, ξ_2 could be slightly larger than the range considered in Lemma 5.9, it is slightly different but the exact same technique applies). Overall, we deduce

$$\mathbb{E} \left[\mathbb{1}_{T_1 = T_2 = 0} \exp \left(i \sum_{k=1}^{M} \frac{\Theta_k \langle X, M_n^* \vec{e}_{j_k} \rangle}{\sqrt{n}} \right) \right]$$

$$= \frac{1}{(2\pi)^2} \int_{-\tau_1}^{\tau_1} \int_{-\tau_2}^{\tau_2} \mathbb{E} \exp \left(i \sum_{k=1}^{M} \frac{\Theta_k \langle \widetilde{X}, M_n^* \vec{e}_{j_k} \rangle}{\sqrt{n}} + i\vec{\xi} \cdot (\widetilde{T}_1, \widetilde{T}_2) \right)$$

$$d\xi_2 d\xi_1 + O(n^{-5/2} (\log n)^{39}),$$

where $\tau_1 = n^{-1/2}(\log n)^8$ and $\tau_2 = n^{-3/2}(\log n)^7$ and $\widetilde{T}_1 = \sum_{j=1}^n \widetilde{X}_j$ and $\widetilde{T}_2 = \sum_{j=1}^n j\widetilde{X}_j$.

Next, use that $\sum_{j=1}^n c_j^* = \sum_{j=1}^n jc_j^* = 0$, which follows from Definition 2.4. This implies that $(\widetilde{T}_1, \widetilde{T}_2)$ is independent of the first part of the sum in the exponential. So, we can factor and integrate over ξ_1, ξ_2 to deduce

$$\mathbb{E}\left[\mathbb{1}_{T_1=T_2=0} \exp\left(i \sum_{k=1}^M \frac{\Theta_k \langle X, M_n^* \vec{e}_{j_k} \rangle}{\sqrt{n}} \right) \right] = q\mathbb{E} \exp\left(i \sum_{k=1}^M \frac{\Theta_k \langle \widetilde{X}, M_n^* \vec{e}_{j_k} \rangle}{\sqrt{n}} \right)$$
$$+ O(n^{-5/2}(\log n)^{39}),$$

where $q = (2\pi)^{-2} \int_{\mathbb{R}^2} \mathbb{E} \exp(i\vec{\xi} \cdot (\widetilde{T}_1, \widetilde{T}_2)) d\vec{\xi}$ (note completing the integral to ∞ does not change the error term). Comparing with the proof of Lemma 9.1, we easily deduce that $q = (1 + O(n^{-1/2}(\log n)^{34}))\mathbb{P}[T_1 = 0 \wedge T_2 = 0]$ and thus also $q = \Omega(n^{-2})$. Dividing by $\mathbb{P}[T_1 = 0 \wedge T_2 = 0]$ and subtracting, we deduce the desired result. □

We deduce an appropriate bound on a moment of $n^{-1/2}|\langle \vec{x}, M_n^* \vec{e}_j \rangle|$.

Lemma 9.5 *Let X_j, T_1, T_2 be as in Definition 5.2 and let $X = (X_1, \ldots, X_n)$. We have*

$$\mathbb{E}\left[\sum_{j=1}^n \left(\frac{|\langle X, M_n^* \vec{e}_j \rangle|}{\sqrt{n}} \right)^3 \Big| T_1 = T_2 = 0 \right]$$
$$\leq C_{9.5} n.$$

Similarly if X_j' is as in Lemma 5.9, we have

$$\mathbb{E}\left[\sum_{j=1}^n \left(\frac{|\langle X', M_n^* \vec{e}_j \rangle|}{\sqrt{n}} \right)^3 \right] \leq C_{9.5} n.$$

Proof The second estimate is trivial by linearity of expectation and Lemma 2.6 (**M3**). For the first estimate, using linearity of expectation it suffices to show $\mathbb{E}|\langle X, M_n^* \vec{e}_j \rangle|^3 = O(n^{3/2})$ uniformly for all $1 \leq j \leq n$.

Note by Bernstein's inequality (Theorem 4.1) that

$$\mathbb{P}[|\langle X, M_n^* \vec{e}_j \rangle| \geq \sqrt{n \log n} \log \log n] \leq n^{-\omega(1)}.$$

As $\mathbb{P}[T_1 = T_2 = 0] = \Omega(n^{-2})$ from Lemma 9.1, we have that

$$\mathbb{P}[|\langle X, M_n^* \vec{e}_j \rangle| \geq \sqrt{n \log n} \log \log n | T_1 = T_2 = 0]$$
$$\leq \frac{\mathbb{P}[|\langle X, M_n^* \vec{e}_j \rangle| \geq \sqrt{n \log n} \log \log n]}{\mathbb{P}[T_1 = T_2 = 0]} \leq n^{-\omega(1)}.$$

Similarly we have that

$$
\mathbb{E}[|\langle X, M_n^* \vec{e}_j \rangle|^3 \mathbb{1}_{|\langle X, M_n^* \vec{e}_j \rangle| \geq \sqrt{n \log n} \log \log n} | T_1 = T_2 = 0]
$$

$$
\leq \frac{\mathbb{E}[|\langle X, M_n^* \vec{e}_j \rangle|^3 \mathbb{1}_{|\langle X, M_n^* \vec{e}_j \rangle| \geq \sqrt{n \log n} \log \log n}]}{\mathbb{P}[T_1 = T_2 = 0]}
$$

$$
\lesssim n^2 \mathbb{E}[|\langle X, M_n^* \vec{e}_j \rangle|^3 \mathbb{1}_{|\langle X, M_n^* \vec{e}_j \rangle| \geq \sqrt{n \log n} \log \log n}]
$$

$$
\lesssim n^2 \sum_{j \geq 0} 2^{3j} \mathbb{P}[|\langle X, M_n^* \vec{e}_j \rangle| \geq 2^j \sqrt{n \log n} \log \log n]
$$

$$
\lesssim n^2 \sum_{j \geq 0} 2^{3j} \exp(-\Omega(4^j \log n (\log \log n)^2)) \lesssim n^{-\omega(1)}, \quad (9.3)
$$

where we have again used Bernstein's inequality (Theorem 4.1) when not conditioning on $T_1 = T_2 = 0$. Therefore it will suffice to consider "reasonable" scales for $|\langle X, M_n^* \vec{e}_j \rangle|$.

We now apply Theorem 4.7 with $\langle X, M_n^* \vec{e}_j \rangle / \sqrt{n}$ conditional on $T_1 = T_2 = 0$ and also with $\langle X', M_n^* \vec{e}_j \rangle / \sqrt{n}$, replacing T by $L = (\log n)^{7/4}$. In particular notice that

$$
\sup_{\tau \in \mathbb{R}} |\mathbb{P}[\langle X, M_n^* \vec{e}_j \rangle \leq \tau \sqrt{n} | T_1 = T_2 = 0] - \mathbb{P}[\langle X', M_n^* \vec{e}_j \rangle \leq \tau \sqrt{n}]|
$$

$$
= \sup_{\tau \in \mathbb{R}} |\mathbb{P}[\langle X, M_n^* \vec{e}_j \rangle / \sqrt{n} \leq \tau | T_1 = T_2 = 0] - \mathbb{P}[\langle X', M_n^* \vec{e}_j \rangle / \sqrt{n} \leq \tau]|
$$

$$
\lesssim \int_{-L}^{L} \frac{|\mathbb{E}[\exp(it\langle X, M_n^* \vec{e}_j \rangle / \sqrt{n}) | T_1 = T_2 = 0] - \mathbb{E}[\exp(it\langle X', M_n^* \vec{e}_j \rangle / \sqrt{n})]|}{|t|} dt
$$

$$
+ \sup_{\tau \in \mathbb{R}} \mathbb{P}[|\langle X', M_n^* \vec{e}_j \rangle / \sqrt{n} - \tau| \leq 1/L]
$$

$$
\lesssim \int_{-L}^{L} \frac{\min\{n^{-1/2}(\log n)^{39}, \mathbb{E}[|t| |\langle X, M_n^* \vec{e}_j \rangle| / \sqrt{n} | T_1 = T_2 = 0] + \mathbb{E}[|t| |\langle X', M_n^* \vec{e}_j \rangle| / \sqrt{n}]\}}{|t|} dt + 1/L
$$

$$
\lesssim \int_{-L}^{L} \frac{\min\{n^{-1/2}(\log n)^{39}, |t| n^{1/2}\}}{|t|} dt + 1/L \lesssim n^{-1/2}(\log n)^{42} + 1/L \lesssim 1/L. \quad (9.4)
$$

The bound on $\sup_{\tau \in \mathbb{R}} \mathbb{P}[|\langle X', M_n^* \vec{e}_j \rangle / \sqrt{n} - \tau| \leq 1/L]$ follows since $\langle X', M_n^* \vec{e}_j \rangle$ is a centered Gaussian random variable and

$$
\sup_{\tau \in \mathbb{R}} \mathbb{P}[|\langle X', M_n^* \vec{e}_j \rangle / \sqrt{n} - \tau| \leq 1/L] = \mathbb{P}[|\langle X', M_n^* \vec{e}_j \rangle / \sqrt{n}| \leq 1/L]
$$

$$
\lesssim \frac{1}{L\sqrt{\mathrm{Var}(\langle X', M_n^* \vec{e}_j \rangle / \sqrt{n})}} \lesssim 1/L
$$

where we use $\|M_n^* \vec{e}_j\|_2 \gtrsim \sqrt{n}$ for all j (which is obvious from **M1**). The first bound between Fourier coefficients in the minimum follows from Lemma 9.4 which shows an error of $O(n^{-1/2}(\log n)^{39})$ between the two resulting Fourier coefficients. To justify

the second bound in the minimum, note that $|\exp(it) - 1| \le |t|$ and therefore

$$
\begin{aligned}
\big|\mathbb{E}[\exp(&it\langle X, M_n^* \vec{e}_j\rangle)|T_1 = T_2 = 0] - \mathbb{E}[\exp(it\langle X', M_n^* \vec{e}_j\rangle)]\big| \\
&= \big|\mathbb{E}[\exp(it\langle X, M_n^* \vec{e}_j\rangle) - 1|T_1 = T_2 = 0] - \mathbb{E}[\exp(it\langle X', M_n^* \vec{e}_j\rangle) - 1]\big| \\
&\le \big|\mathbb{E}[|t||\langle X, M_n^* \vec{e}_j\rangle||T_1 = T_2 = 0]\big| + \big|\mathbb{E}[|t||\langle X', M_n^* \vec{e}_j\rangle|]\big| \\
&\le |t|\big(|\mathbb{E}[|\langle X, M_n^* \vec{e}_j\rangle)|^3|T_1 = T_2 = 0]\big)^{1/3} + |t|(|\mathbb{E}[|\langle X', M_n^* \vec{e}_j\rangle)|^3|])^{1/3} \lesssim |t|n.
\end{aligned}
$$

In the final inequality, the first term is controlled using the tail bound (9.3), which allows us to focus on cases where $|\langle X, M_n^* \vec{e}_j\rangle| < \sqrt{n \log n} \log \log n$ (in which case the claimed bound is evident). The second term, coming from a Gaussian model, is easy to control using Lemma 2.6 (**M3**).

Finally, we use (9.4) and apply integration by parts. We have

$$
\begin{aligned}
\mathbb{E}\Big[|\langle X, M_n^* \vec{e}_j\rangle|^3 &\mathbb{1}_{|\langle X, M_n^* \vec{e}_j\rangle| < \sqrt{n \log n} \log \log n}|T_1 = T_2 = 0\Big] \\
&= \mathbb{E}\Big[|\langle X', M_n^* \vec{e}_j\rangle|^3 \mathbb{1}_{|\langle X', M_n^* \vec{e}_j\rangle| < \sqrt{n \log n} \log \log n}\Big] + O((\sqrt{n \log n} \log \log n)^3/L) \\
&= O(n^{3/2}),
\end{aligned}
$$

using that X' is Gaussian (and **M4**) to control the first term in the second line. $\qquad\square$

We will also require the following variant of Lemma 9.5 which bounds the difference between nearby coordinates.

Lemma 9.6 *Let* X_j, T_1, T_2 *be as in Definition* 5.2 *and let* $X = (X_1, \ldots, X_n)$. *If* $\varepsilon \ge 1/\log \log n$, *and* $|j - j'| \le \varepsilon n$ *then we have*

$$
\mathbb{E}\left[\frac{||\langle X, M_n^* \vec{e}_j\rangle|^2 - |\langle X, M_n^* \vec{e}_{j'}\rangle|^2|}{n}\bigg|T_1 = T_2 = 0\right] \le C_{9.5}\varepsilon^{1/2}.
$$

Similarly if X'_j *is as in Lemma* 5.9, *we have*

$$
\mathbb{E}\left[\frac{||\langle X', M_n^* \vec{e}_j\rangle|^2 - |\langle X', M_n^* \vec{e}_{j'}\rangle|^2|}{n}\right] \le C_{9.5}\varepsilon^{1/2}.
$$

Proof First, we may assume that $|j - j'| \ge \varepsilon n/4$ (i.e., the "separated" case). Indeed, given values with $|j - j'| < \varepsilon n/4$ we may choose j^* which is at least at $\varepsilon n/4$ distance from both j, j' (and still in the range $[n]$). Then we can apply the result for the separated case twice on the pairs j, j^* and j^*, j' and then use the triangle inequality.

As $|y^2 - z^2| = |y - z| \cdot |y + z| \le |y - z| \cdot (|y| + |z|)$, by Cauchy–Schwarz it suffices to prove that

$$
\begin{aligned}
\mathbb{E}\left[\frac{|\langle X, M_n^* \vec{e}_j\rangle|^2}{n}\bigg|T_1 = T_2 = 0\right] &\le C_{9.5}, \\
\mathbb{E}\left[\frac{|\langle X, M_n^* \vec{e}_j\rangle - \langle X, M_n^* \vec{e}_{j'}\rangle|^2}{n}\bigg|T_1 = T_2 = 0\right] &\le C_{9.5}\varepsilon
\end{aligned}
\tag{9.5}
$$

and analogous estimates for X'. The estimates (9.5) follow immediately for X' by Lemma 2.6 (in particular **M7** for the second estimate). For X, note that by Hölder's inequality a strictly stronger estimate than the first estimate in (9.5) is given in the proof of Lemma 9.5 (which shows $\mathbb{E}|\langle X, M_n^* \vec{e}_j \rangle|^3 = O(n^{3/2})$ uniformly in j). We now focus on the second estimate of (9.5).

We apply Theorem 4.7 with $\langle X, M_n^*(\vec{e}_j - \vec{e}_{j'})\rangle/\sqrt{n}$ conditional on $T_1 = T_2 = 0$ and also with $\langle X', M_n^*(\vec{e}_j - \vec{e}_{j'})\rangle/\sqrt{n}$, replacing T by $L = (\log n)^{7/4}$. In particular notice that

$$
\sup_{\tau \in \mathbb{R}} |\mathbb{P}[\langle X, M_n^*(\vec{e}_j - \vec{e}_{j'})\rangle/\sqrt{n} \leq \tau | T_1 = T_2 = 0] - \mathbb{P}[\langle X', M_n^*(\vec{e}_j - \vec{e}_{j'})\rangle/\sqrt{n} \leq \tau]|
$$

$$
\lesssim \int_{-L}^{L} \frac{|\mathbb{E}[\exp(it\langle X, M_n^*(\vec{e}_j - \vec{e}_{j'})\rangle)/\sqrt{n})|T_1 = T_2 = 0] - \mathbb{E}[\exp(it\langle X', M_n^*(\vec{e}_j - \vec{e}_{j'})\rangle)/\sqrt{n})]|}{|t|} dt + \varepsilon^{-1}/L
$$

$$
\lesssim \int_{-L}^{L} \frac{\min\{n^{-1/2}(\log n)^{39}, \mathbb{E}[|t||\langle X, M_n^*(\vec{e}_j - \vec{e}_{j'})\rangle||T_1 = T_2 = 0] + \mathbb{E}[|t||\langle X', M_n^*(\vec{e}_j - \vec{e}_{j'})\rangle|]\}}{|t|} dt + \varepsilon^{-1}/L
$$

$$
\lesssim \int_{-L}^{L} \frac{\min\{n^{-1/2}(\log n)^{39}, |t|n^4\}}{|t|} dt + 1/L \lesssim n^{-1/2}(\log n)^{42} + \varepsilon^{-1}/L \lesssim \varepsilon^{-1}/L. \tag{9.6}
$$

The error term ε^{-1}/L is deduced in a similar manner to the $1/L$ term in (9.4), except that instead we have $\|M_n^*(\vec{e}_j - \vec{e}_{j'})\|_2 \gtrsim \varepsilon\sqrt{n}$ due to $|j - j'| \in [\varepsilon n/4, \varepsilon n]$ and **M1**. The first bound between Fourier coefficients in the minimum follows from Lemma 9.4 (with $M = 2$) which shows an error of $O(n^{-1/2}(\log n)^{39})$ between the two resulting Fourier coefficients. To justify the second bound in the minimum, note that $|\exp(it) - 1| \leq |t|$ and therefore

$$
\big|\mathbb{E}[\exp(it\langle X, M_n^*(\vec{e}_j - \vec{e}_{j'}))\rangle)|T_1 = T_2 = 0] - \mathbb{E}[\exp(it\langle X', M_n^*(\vec{e}_j - \vec{e}_{j'}))\rangle)]\big|
$$

$$
= \big|\mathbb{E}[\exp(it\langle X, M_n^*(\vec{e}_j - \vec{e}_{j'}))\rangle) - 1|T_1 = T_2 = 0] - \mathbb{E}[\exp(it\langle X', M_n^*(\vec{e}_j - \vec{e}_{j'}))\rangle) - 1]\big|
$$

$$
\leq \big|\mathbb{E}[|t||\langle X, M_n^*(\vec{e}_j - \vec{e}_{j'}))\rangle||T_1 = T_2 = 0]\big| + \big|\mathbb{E}[|t||\langle X', M_n^*(\vec{e}_j - \vec{e}_{j'})\rangle|]\big|
$$

$$
\leq |t|\mathbb{P}[T_1 = T_2 = 0]^{-1}\big|\mathbb{E}[|\langle X, M_n^*(\vec{e}_j - \vec{e}_{j'}))\rangle|]\big| + |t|n \lesssim |t|n^4.
$$

In the final inequality, to control the first term we use $\mathbb{P}[T_1 = T_2 = 0] = \Omega(n^{-2})$ and then we use Bernstein's inequality to control the moment.

Furthermore, by Bernstein's inequality (Theorem 4.1) and a similar argument as in the derivation of (9.3) we have

$$
\mathbb{E}\left[|\langle X, M_n^*(\vec{e}_j - \vec{e}_{j'})\rangle|^2 \mathbb{1}_{|\langle X, M_n^*(\vec{e}_j - \vec{e}_{j'})\rangle| \geq \sqrt{n \log n}(\log \log n)}\right] = n^{-\omega(1)}.
$$

Since $\mathbb{P}[T_1 = T_2 = 0] = \Omega(n^{-2})$ we deduce

$$
mbm E\left[|\langle X, M_n^*(\vec{e}_j - \vec{e}_{j'})\rangle|^2 \mathbb{1}_{|\langle X, M_n^*(\vec{e}_j - \vec{e}_{j'})\rangle| \geq \sqrt{n \log n}(\log \log n)}|T_1 = T_2 = 0\right]
$$

$$
\leq \frac{n^{-\omega(1)}}{\mathbb{P}[T_1 = T_2 = 0]} = n^{-\omega(1)}.
$$

Finally, by (9.6) and integration by parts we have

$$
\begin{aligned}
&\mathbb{E}\Big[|\langle X, M_n^*(\vec{e}_j - \vec{e}_{j'})\rangle|^2 \mathbb{1}_{|\langle X, M_n^*(\vec{e}_j - \vec{e}_{j'})\rangle| < \sqrt{n \log n} \log \log n} \Big| T_1 = T_2 = 0\Big] \\
&= \mathbb{E}\Big[|\langle X', M_n^*(\vec{e}_j - \vec{e}_{j'})\rangle|^2 \mathbb{1}_{|\langle X', M_n^*(\vec{e}_j - \vec{e}_{j'})\rangle| < \sqrt{n \log n} \log \log n}\Big] \\
&\quad + O((\sqrt{n \log n} \log \log n)^2 \varepsilon^{-1}/L) \\
&\lesssim \varepsilon n + O(n (\log \log n)^3 (\log n)^{-3/4}),
\end{aligned}
$$

The last line follows from the result for the Gaussian model X' deduced above (which relies on the assumption $|j - j'| \le \varepsilon n$). $\qquad\square$

We next deduce that $\min_{a,b \in \mathbb{R}} \sum_j (c_j - a - bj)^2$ satisfies an appropriate anticoncentration bound near 0 with high probability, so as to control singularity behavior.

Lemma 9.7 *Let X_j, T_1, T_2 be as in Definition 5.2 and let $\varepsilon \ge 1/\log n$. Let $X = (X_1, \ldots, X_n)$. We have*

$$
\mathbb{P}\Big[\sum_{j=1}^n \langle X, M_n^* \vec{e}_j\rangle^2 \le \varepsilon n^2 \Big| T_1 = T_2 = 0\Big] \lesssim \varepsilon^4.
$$

Proof Let k be a sufficiently large absolute integer constant. For $1 \le t \le k$ define $S_t = [tn/(k+1), tn/(k+1) + \delta^2 n]$, $S_t' = [tn/(k+1) + 2\delta^2 n, tn/(k+1) + 3\delta^2 n]$, where $\delta \in (0, 1/2)$ will be a sufficiently small constant (with respect to k) to be chosen later. Let

$$
\mathcal{I} = \{(j_1, j_1', j_2, j_2', \ldots, j_k, j_k')| : j_t \in S_t, j_t' \in S_t', |\langle X, M_n^*(\vec{e}_{j_t'} - \vec{e}_{j_t})\rangle| \le \delta^{-2} \varepsilon^{1/2} n^{1/2}\}.
$$

If $\sum_{j=1}^n \langle X, M_n^* \vec{e}_j\rangle^2 \le \varepsilon n^2$, by Markov's inequality there are fewer than $\delta^4 n$ indices j such that $|\langle X, M_n^* \vec{e}_j\rangle| \ge \delta^{-2} \varepsilon^{1/2} n^{1/2}$. Therefore it follows that $|\mathcal{I}| \ge (\delta^2 n/2)^{2k}$ under this event.

We now assume that for any choices of $j_t \in S_t$, $j_t' \in S_t'$ for $1 \le t \le k$ that

$$
\mathbb{P}\Big[\bigcap_{t=1}^k |\langle X, M_n^*(\vec{e}_{j_t'} - \vec{e}_{j_t})\rangle| \le \delta^{-2} \varepsilon^{1/2} n^{1/2} \Big| T_1 = T_2 = 0\Big] \lesssim_{k,\delta} \varepsilon^{k/2}; \quad (9.7)
$$

we will justify this at the end of the proof. Given (9.7), we compute that

$$
\begin{aligned}
\mathbb{E}[|\mathcal{I}||T_1 = T_2 = 0] &\le \sum_{\substack{j_t \in S_t, j_t' \in S_t' \\ \forall t \in [k]}} \mathbb{P}\Big[\bigcap_{t=1}^k |\langle X, M_n^*(\vec{e}_{j_t'} - \vec{e}_{j_t})\rangle| \le \delta^{-2} \varepsilon^{1/2} n^{1/2} \Big| T_1 = T_2 = 0\Big] \\
&\lesssim_{k,\delta} \varepsilon^{k/2} n^{2k}. \qquad\qquad (9.8)
\end{aligned}
$$

Note (9.8) immediately implies the desired result taking any integer $k \geq 8$ and applying Markov's inequality: we find that $|\mathcal{I}| \geq (\delta^2 n/2)^{2k}$ occurs with probability $O(\varepsilon^4)$, which implies the same for the original event by the earlier analysis.

To prove (9.7) the idea is to use Theorem 4.8 and Lemma 9.4. Writing B for the radius $k\delta^2/\varepsilon^{1/2}$ unit ball in $\mathbb{R}^k$, we have

$$\mathbb{P}\left[\bigcap_{t=1}^{k} \frac{|\langle X, M_n^*(\vec{e}_{j_t'} - \vec{e}_{j_t})\rangle|}{\sqrt{n}} \leq \delta^{-2}\varepsilon^{1/2} \,\Big|\, T_1 = T_2 = 0\right]$$

$$\lesssim_{\delta,k} \varepsilon^{k/2} \int_B \left|\mathbb{E}\left[\exp\left(2\pi i \sum_{t=1}^{k} \frac{\xi_t \langle X, M_n^*(\vec{e}_{j_t'} - \vec{e}_{j_t})\rangle}{\sqrt{n}}\right) \,\Big|\, T_1 = T_2 = 0\right]\right| d\vec{\xi}$$

$$\lesssim_{\delta,k} \varepsilon^{k/2} \int_B \left|\mathbb{E}\exp\left(2\pi i \sum_{t=1}^{k} \frac{\xi_t \langle \widetilde{X}, M_n^*(\vec{e}_{j_t'} - \vec{e}_{j_t})\rangle}{\sqrt{n}}\right)\right| d\vec{\xi} + O(n^{-1/2}(\log n)^{39})$$

$$\lesssim_{\delta,k} \varepsilon^{k/2} q + O(n^{-1/2}(\log n)^{39}),$$

where q is the probability density function of the Gaussian vector $(n^{-1/2}\langle \widetilde{X}, M_n^*(\vec{e}_{j_t'} - \vec{e}_{j_t})\rangle)_{1 \leq t \leq k}$ evaluated at 0. For the last line, we used the nonnegativity of Gaussian characteristic functions and Fourier inversion. Now we show $q = O_{\delta,k}(1)$ to finish.

Note that by **M1** and explicit computation we have

$$M_n^*(\vec{e}_{j_t'} - \vec{e}_{j_t}) = \frac{1}{2} \sum_{j_t < r < j_t'} \vec{e}_r + \vec{v}_{j_t, j_t'},$$

where $\|\vec{v}_{j_t, j_t'}\|_2 \lesssim \delta^2 n^{1/2}$ and note that $\|\sum_{j_t < r < j_t'} \vec{e}_r\|_2 \gtrsim \delta n^{1/2}$. Therefore we find that

$$\text{dist}(M_n^*(\vec{e}_{j_t'} - \vec{e}_{j_t}), \text{span}_{\mathbb{R}}\{M_n^*(\vec{e}_{j_s'} - \vec{e}_{j_s}): s \in [k] \setminus \{t\}\}) \geq \|M_n^*(\vec{e}_{j_t'} - \vec{e}_{j_t})\|_2/2$$

if δ is sufficiently small as a function of k. This implies that the covariance matrix of the above Gaussian vector is diagonally dominated with constant order diagonal entries, and the result follows. $\quad\square$

We now conclude with the proof of Theorem 1.8.

Proof of Theorem 1.8 Given a die B with the frequency counts $(\widetilde{b}_j)_{1 \leq j \leq n}$, let $y_j = \sum_{1 \leq k < j} \widetilde{b}_k + \widetilde{b}_j/2 - (j - 1/2)$. We say B is *suitable* if $(y_j)_{1 \leq j \leq n}$ is coarse (Definition 5.1). By Lemma 7.1 we have that B is coarse with probability $1 - n^{-\omega(1)}$. Furthermore, write $\widetilde{b} = (\widetilde{b}_1, \ldots, \widetilde{b}_n)$ and note

$$\min_{a,b \in \mathbb{R}} \sum_{j=1}^{n} (y_j - aj - b)^2 = \|M_n^* \widetilde{b}\|_2^2, \tag{9.9}$$

which can be seen by looking at projections of vectors in $\mathbb{R}^n$, and also noting that $\sum_{j=1}^{n}(\widetilde{b}_j - 1) = \sum_{j=1}^{n} j(\widetilde{b}_j - 1) = 0$.

Recall that by Lemma 2.2, $(\widetilde{b}_1, \ldots, \widetilde{b}_n)$ has the distribution of $(X_1, \ldots, X_n)$ conditional on $T_1 = T_2 = 0$, borrowing the setup of Definition 5.2 (where $\Delta = \text{Geom}(1/2)$ for the multiset model and $\Delta = \text{Pois}(1)$ for the balanced sequence model).

Therefore, applying Lemma 9.1 we find

$$
\begin{aligned}
\mathbb{P}[A \text{ ties } B] &= \mathbb{P}[A \text{ ties } B \wedge B \text{ is suitable}] \pm n^{-\omega(1)} \\
&= \mathbb{E}_B\big[\mathbb{P}[A \text{ ties } B | B \text{ is suitable}]\big] \pm n^{-\omega(1)} \\
&= \mathbb{E}\left[\frac{\mathbb{1}_{B \text{ is suitable}}}{(8\pi \text{Var}[\Delta])^{1/2}\|M_n^*\widetilde{b}\|_2}\right] \pm n^{-3/2}(\log n)^{40}.
\end{aligned}
$$

The final line follows from using Lemma 2.2 to interpret the probability of a tie as the ratio of the two expressions in Lemma 9.1, and using **S4** to control the resulting error terms.

Fix a constant $\varepsilon > 0$. We will take $\varepsilon \to 0^+$ sufficiently slowly at the end of the proof. We have

$$
\begin{aligned}
\mathbb{E}\left[\frac{\mathbb{1}_{B \text{ is suitable}}}{\|M_n^*\widetilde{b}\|_2} - \frac{\mathbb{1}_{B \text{ is suitable}}}{\max(\|M_n^*\widetilde{b}\|_2, \varepsilon n)}\right] &\lesssim n^{-\omega(1)} + \sum_{(\log n)^{-2} \leq 2^{-j} \leq \varepsilon} \frac{2^j}{n} \cdot \mathbb{P}[\|M_n^*\widetilde{b}\|_2 \leq 2^{-j}n] \\
&\lesssim n^{-\omega(1)} + \sum_{(\log n)^{-2} \leq 2^{-j} \leq \varepsilon} \frac{2^j}{n} \cdot 4^{-j} \\
&\lesssim \frac{\varepsilon}{n},
\end{aligned}
$$

where we have dyadically decomposed the small values of $\|M_n^*\widetilde{b}\|_2$ and applied Lemma 9.7 (we apply the lemma for $\max(2^{-j}, (\log n)^{-1})$). Note that $\langle X, M_n^*\vec{e}_j\rangle = -\langle M_n^*X, \vec{e}_j\rangle$ is the negative of the jth coordinate of M_n^*X by **M2**, and we again used Lemma 2.2. Additionally, we are using that $\|M_n^*\vec{b}\|_2 \geq n(\log n)^{-2}$ for coarse $\vec{b}$ by (9.9) and **S4**.

Therefore

$$
\begin{aligned}
(8\pi \text{Var}[\Delta])^{1/2}\mathbb{P}[A \text{ ties } B] &= \mathbb{E}\left[\frac{\mathbb{1}_{B \text{ is suitable}}}{\max(\|M_n^*\widetilde{b}\|_2, \varepsilon n)}\right] \pm O(\varepsilon n^{-1}) \\
&= \mathbb{E}\left[\frac{1}{\max(\|M_n^*\widetilde{b}\|_2, \varepsilon n)}\right] \pm O(\varepsilon n^{-1})
\end{aligned}
$$

provided that n is sufficiently large with respect to ε.

The next idea is to approximate $\|M_n^*\widetilde{b}\|_2$ via sampling random coordinates $j_1, \ldots, j_T$ for a sufficiently large value of T and then estimating the L^2-norm of the vector $M_n^*\widetilde{b}$ via examining only these coordinates. This converts understanding a quadratic form into a question of purely linear forms. Let $\mathcal{E}_1$ denote the event that $\|M_n^*\widetilde{b}\|_3/n^{1/3} \leq \varepsilon^{-1}n^{1/2}$. Lemma 9.5 and Markov's inequality implies that $\mathcal{E}_1$ holds with probability at least $1 - \varepsilon^3$, hence

$$\mathbb{E}\left[\frac{1}{\max(\|M_n^*\widetilde{b}\|_2, \varepsilon n)}\right] = \mathbb{E}\left[\frac{\mathbb{1}_{\mathcal{E}_1}}{\max(\|M_n^*\widetilde{b}\|_2, \varepsilon n)}\right] \pm O(\varepsilon^2 n^{-1}). \qquad (9.10)$$

Let $\psi \colon \mathbb{R} \to \mathbb{R}_{\geq 0}$ be smooth with $\psi(y) = 0$ for $y \leq \varepsilon/2$ and $\psi(y) = y$ for $y \geq \varepsilon$, $0 \leq \psi(y) \leq y$ for $y \in [\varepsilon/2, \varepsilon]$, and such that ψ is $O(1/\varepsilon)$-Lipschitz (a construction can be derived in a standard manner using bump functions). Let $g(x) := \psi(1/\max(x, \varepsilon))$. Then

$$(8\pi \operatorname{Var}[\Delta])^{1/2} n \mathbb{P}[A \text{ ties } B] = \mathbb{E}[\mathbb{1}_{\mathcal{E}_1} g(\|M_n^*\widetilde{b}\|_2/n)] \pm O(\varepsilon) \qquad (9.11)$$

since $\mathcal{E}_1$ implies $\|M_n^*\widetilde{b}\|_2/n^{1/2} \leq \varepsilon^{-1} n^{1/2}$.

Now we use Lemma 9.3. Given the event $\mathcal{E}_1$ we have $\|M_n^*\widetilde{b}\|_{3/2}/n^{2/3} \leq \|M_n^*\widetilde{b}\|_3/n^{1/3} \leq \varepsilon^{-1} n^{1/2}$. Thus a uniformly random sample of $T = \lceil \varepsilon^{-100} \rceil$ independent coordinates $j_1, \ldots, j_T$ satisfies

$$\left|\frac{1}{n}\|M_n^*\widetilde{b}\|_2^2 - \frac{1}{T}\sum_{1 \leq k \leq T}\langle \vec{e}_{j_k}, M_n^*\widetilde{b}\rangle^2\right| \leq \varepsilon^{20} n \qquad (9.12)$$

with probability at least $1 - \varepsilon^{10}$ in such a situation.

Now consider the function $h(x) := g(x^{1/2})$. Due to the properties of ψ, h is $O(\varepsilon^{-4})$-Lipschitz since $x \mapsto 1/\max(x^{1/2}, \varepsilon)$ is $O(\varepsilon^{-3})$-Lipschitz and ψ is $O(1/\varepsilon)$-Lipschitz. Now we can use this and (9.12) to immediately deduce

$$\left|\mathbb{E}[\mathbb{1}_{\mathcal{E}_1} g(\|M_n^*\widetilde{b}\|_2/n)] - \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{\mathcal{E}_1} g\left(\left(\frac{1}{nT}\sum_{k=1}^T \langle M_n^*\widetilde{b}, \vec{e}_{j_k}\rangle^2\right)^{1/2}\right)\bigg| j_1, \ldots, j_T\right]\right]\right.$$
$$\left. \leq O(\varepsilon^{-4}) \cdot \varepsilon^{20} + \varepsilon^{10},\right.$$

where the second term encapsulates the probability of failure of the event (9.12). Since $\mathbb{P}[\mathcal{E}_1] \geq 1 - \varepsilon^3$ and g is clearly bounded by $1/\varepsilon$, it follows immediately that

$$\left|\mathbb{E}[\mathbb{1}_{\mathcal{E}_1} g(\|M_n^*\widetilde{b}\|_2/n)] - \mathbb{E}\left[\mathbb{E}\left[g\left(\left(\frac{1}{nT}\sum_{k=1}^T \langle M_n^*\widetilde{b}, \vec{e}_{j_k}\rangle^2\right)^{1/2}\right)\bigg| j_1, \ldots, j_T\right]\right]\right.$$
$$\left. \lesssim \varepsilon^{16} + (1/\varepsilon) \cdot \varepsilon^3. \right. \qquad (9.13)$$

Now let $L = \lceil \varepsilon^{-1000} \rceil$ and $n' = \lfloor \varepsilon^{1000} n \rfloor$. For each k let $\widetilde{j}_k$ denote the nearest index to j_k in the set $\{n', 2n', \ldots, Ln'\}$. Recall that h is $O(\varepsilon^{-4})$-Lipschitz. So, applying Lemma 9.6, we have

$$\left| \mathbb{E}\left[\mathbb{E}\left[g\left(\left(\frac{1}{nT} \sum_{k=1}^{T} \langle M_n^* \widetilde{b}, \vec{e}_{j_k} \rangle^2 \right)^{1/2} \right) \middle| j_1, \ldots, j_T \right] \right] \right.$$

$$\left. - \mathbb{E}\left[\mathbb{E}\left[g\left(\left(\frac{1}{nT} \sum_{k=1}^{T} \langle M_n^* \widetilde{b}, \vec{e}_{\widetilde{j}_k} \rangle^2 \right)^{1/2} \right) \middle| j_1, \ldots, j_T \right] \right] \right|$$

$$\lesssim \varepsilon^{-4} \mathbb{E}\left[\mathbb{E}\left[\left| \frac{1}{nT} \sum_{k=1}^{T} \langle M_n^* \widetilde{b}, \vec{e}_{j_k} \rangle^2 - \frac{1}{nT} \sum_{k=1}^{T} \langle M_n^* \widetilde{b}, \vec{e}_{\widetilde{j}_k} \rangle^2 \right| \middle| j_1, \ldots, j_T \right] \right]$$

$$\lesssim \varepsilon^{-4} (\varepsilon^{1000})^{1/2} = \varepsilon^{496}.$$

As noted at the beginning of this proof, Lemma 2.2 implies that $\widetilde{b} = (\widetilde{b}_1, \ldots, \widetilde{b}_n)$ has the distribution of $(X_1, \ldots, X_n)$ conditional on $T_1 = T_2 = 0$ with X defined as in Definition 5.2. Now by Lévy continuity, Lemma 9.4, and **M1** we see that the distributions of

$$\left(\frac{\langle M_n^* \widetilde{b}, \vec{e}_{kn'} \rangle}{\sqrt{n}} \right)_{1 \le k \le L}, \qquad \left(\frac{\langle M_n^* \widetilde{X}, \vec{e}_{kn'} \rangle}{\sqrt{n}} \right)_{1 \le k \le L}$$

each converge jointly to the same fixed distribution which is independent of n (but depending on ε). As g is a bounded and continuous function, for n sufficiently large by the Portmanteau theorem we deduce that

$$\left| \mathbb{E}\left[\mathbb{E}\left[g\left(\left(\frac{1}{nT} \sum_{k=1}^{T} \langle M_n^* \widetilde{b}, \vec{e}_{\widetilde{j}_k} \rangle^2 \right)^{1/2} \right) \middle| j_1, \ldots, j_T \right] \right] \right.$$

$$\left. - \mathbb{E}\left[\mathbb{E}\left[g\left(\left(\frac{1}{nT} \sum_{k=1}^{T} \langle M_n^* \widetilde{X}, \vec{e}_{\widetilde{j}_k} \rangle^2 \right)^{1/2} \right) \middle| j_1, \ldots, j_T \right] \right] \right| \le \varepsilon^2, \qquad (9.14)$$

say. Finally, using Lemmas 9.3, 9.5 and 9.6 in the Gaussian model instead and mimicking the above argument (for (9.13)) in reverse demonstrates

$$\left| \mathbb{E}[\mathbb{1}_{\mathcal{E}_2} g(\|M_n^* \widetilde{X}\|_2 / n)] - \mathbb{E}\left[\mathbb{E}\left[g\left(\left(\frac{1}{nT} \sum_{k=1}^{T} \langle M_n^* \widetilde{X}, \vec{e}_{\widetilde{j}_k} \rangle^2 \right)^{1/2} \right) \middle| j_1, \ldots, j_T \right] \right] \right| \lesssim \varepsilon^2,$$
$$(9.15)$$

where $\mathcal{E}_2$ is the event that $\|M_n^* \widetilde{X}\|_3 / n^{1/3} \le \varepsilon^{-1} n^{1/2}$. Combining (9.11) and (9.13)–(9.15), we deduce

$$(8\pi \operatorname{Var}[\Delta])^{1/2} n \mathbb{P}[A \text{ ties } B] = \mathbb{E}[\mathbb{1}_{\mathcal{E}_2} g(\|M_n^* \widetilde{X}\|_2 / n)] \pm O(\varepsilon).$$

Finally, we note $\mathbb{P}[\mathcal{E}_2] \ge 1 - \varepsilon^3$ similar to the argument for $\mathcal{E}_1$ (namely, Lemma 9.5 and Markov's inequality); we have already implicitly used this when mimicking the argument in the Gaussian model. Since g is bounded by $1/\varepsilon$, we deduce

$$(8\pi\,\mathrm{Var}[\Delta])^{1/2}n\mathbb{P}[A \text{ ties } B] = \mathbb{E}[g(\|M_n^*\widetilde{X}\|_2/n)] \pm O(\varepsilon). \tag{9.16}$$

Finally, we study the distribution of $\|M_n^*\widetilde{X}\|_2/n$. Recall from Definition 2.4 that the spectrum of M_n^* (perhaps minus one copy of 0) is $\{\pm i\sigma_{n,\ell} : \ell \in [\lfloor n/2 \rfloor]\}$ and M_n^* is skew-symmetric. Thus the eigenvalues of the symmetric matrix $(M_n^*)^\mathsf{T} M_n^*$ are composed of two copies of $\sigma_{n,\ell}^2$ (with corresponding eigenvectors $\vec{u}_\ell, \vec{u}_\ell'$) for $\ell \in [\lfloor n/2 \rfloor]$ and maybe 0. We choose these eigenvectors so that they are orthonormal. Therefore, by the spectral theorem,

$$\|M_n^*\widetilde{X}\|_2^2 = \widetilde{X}^\mathsf{T}(M_n^*)^\mathsf{T} M_n^* \widetilde{X} = \sum_{\ell=1}^{\lfloor n/2 \rfloor} (\sigma_{n,\ell}^2 \langle \widetilde{X}, \vec{u}_\ell \rangle^2 + \sigma_{n,\ell}^2 \langle \widetilde{X}, \vec{u}_\ell' \rangle^2).$$

Finally, letting $W_\ell, W_\ell' \sim \mathcal{N}(0, \mathrm{Var}[\Delta])$ for $\ell \geq 1$ (such that all these variables are mutually independent) and using orthogonal invariance of Gaussian vectors, this can be rewritten as

$$\frac{\|M_n^*\widetilde{X}\|_2}{n} \stackrel{d.}{=} \left(\sum_{\ell=1}^{\lfloor n/2 \rfloor} \frac{\sigma_{n,\ell}^2}{n^2}(W_\ell^2 + W_\ell'^2) \right)^{1/2}.$$

Now by **M6** and **M8** and Theorem 4.6 we deduce

$$\mathbb{E}[g(\|M_n^*\widetilde{X}\|_2/n)] = \mathbb{E}\left[g\left(\left(\sum_{\ell=1}^{\lfloor n/2 \rfloor} \frac{\sigma_{n,\ell}^2}{n^2}(W_\ell^2 + W_\ell'^2) \right)^{1/2} \right) \right]$$

$$= \mathbb{E}\left[g\left(\left(\sum_{\ell=1}^{\lfloor \varepsilon^{-8} \rfloor} \frac{\sigma_{n,\ell}^2}{n^2}(W_\ell^2 + W_\ell'^2) \right)^{1/2} \right) \right] \pm O(\varepsilon^2)$$

$$= \mathbb{E}\left[g\left(\left(\sum_{\ell=1}^{\lfloor \varepsilon^{-8} \rfloor} \sigma_\ell^2(W_\ell^2 + W_\ell'^2) \right)^{1/2} \right) \right] \pm O(\varepsilon^2)$$

as long as n is large in terms of ε. We have used the Lipschitz property of h here. We claim that taking the limit $\varepsilon \to 0^+$ gives the result. To check this, we note that for all $\rho \geq 0$ we have

$$\mathbb{P}\left[\sum_{\ell \geq 1} \sigma_\ell^2(W_\ell^2 + W_\ell'^2) \leq \rho \right] \leq \mathbb{P}\left[\bigcap_{1 \leq \ell \leq 5} \sigma_\ell^2(W_\ell^2 + W_\ell'^2) \leq \rho \right]$$

$$\leq \prod_{\ell=1}^{5} (\mathbb{P}[W_\ell^2 \leq K\rho]\mathbb{P}[W_\ell'^2 \leq K\rho]) \lesssim \rho^5$$

for an appropriate absolute constant K. Additionally,

$$\mathbb{P}\left[\sum_{\ell \geq 1} \sigma_\ell^2 (W_\ell^2 + W_\ell'^2) \geq \varepsilon^{-1}\right] \leq \exp(-\varepsilon^{-\Omega(1)})$$

by Theorem 4.6. Therefore, we can absorb the difference between $g(y)$ and $1/y$ without any issue, uniformly in the limit. That is,

$$\lim_{\varepsilon \to 0^+} \mathbb{E}\left[g\left(\left(\sum_{\ell=1}^{\lfloor \varepsilon^{-8} \rfloor} \sigma_\ell^2 (W_\ell^2 + W_\ell'^2) \right)^{1/2} \right) \right] = \mathbb{E}\left[\left(\sum_{\ell \geq 1} \sigma_\ell^2 (W_\ell^2 + W_\ell'^2) \right)^{-1/2} \right]$$

(note that g depends on ε here). Combining these final equalities with (9.16) and taking ε to go slowly to 0, and letting $(W_\ell, W_\ell') = \sqrt{\mathrm{Var}[\Delta]}(Z_\ell, Z_\ell')$ for standard Gaussians Z_ℓ, Z_ℓ', we ultimately deduce

$$(8\pi \mathrm{Var}[\Delta])^{1/2} n \mathbb{P}[A \text{ ties } B] = (\mathrm{Var}[\Delta])^{-1/2} \mathbb{E}\left[\left(\sum_{\ell \geq 1} \sigma_\ell^2 (Z_\ell^2 + Z_\ell'^2) \right)^{-1/2} \right] + o(1).$$

Rearranging, this agrees with the desired Theorem 1.8. □

We end by briefly discussing an (amusing) interpretation of the constant α corresponding to ordinary least squares regression in the context of Brownian motion. The above proof implicitly shows that given a fixed set of indices $j_1, \ldots, j_k$, $(\vec{e}_{j_t} M_n^* X)_{1 \leq t \leq k}$ in distribution converges toward a snapshot of a Brownian motion at the times j_t/n, where the Brownian motion is conditioned to end at 0 at time 1 and conditioned to have total signed area under the Brownian motion equal to 0. Note that then $\min_{a,b \in \mathbb{R}} (c_j - a - bj)^2$ corresponds to approximating such a Brownian motion by the best linear-function fit coming from Ordinary Least Squares regression. We leave making this precise an exercise for the reader.

Declarations

References

1. Akin, E.: Generalized intransitive dice: mimicking an arbitrary tournament. J. Dyn. Games **8**, 1–20 (2021)
2. Akin, E., Saccamano, J.: Generalized intransitive dice II: partition constructions. J. Dyn. Games **8**, 187–202 (2021)
3. Angel, L., Davis, M.: A direct construction of nontransitive dice sets. J. Comb. Des. **25**, 523–529 (2017)
4. Berkowitz, R.: A quantitative local limit theorem for triangles in random graphs. arXiv:1610.01281
5. Carbery, A., Wright, J.: Distributional and L^q norm inequalities for polynomials over convex bodies in $\mathbb{R}^n$. Math. Res. Lett. **8**, 233–248 (2001)
6. Chung, F.R.K., Graham, R.L.: Quasi-random tournaments. J. Graph Theory **15**, 173–198 (1991)
7. Conrey, B., Gabbard, J., Grant, K., Liu, A., Morrison, K.E.: Intransitive dice. Math. Mag. **89**, 133–143 (2016)
8. Cornacchia, E., Hązła, J.: Intransitive dice tournament is not quasirandom. arXiv:2011.10067
9. Diaconis, P., Janson, S.: Graph limits and exchangeable random graphs. Rend. Mat. Appl. **7**(28), 33–61 (2008)
10. Finkelstein, M, Thorp, E.O.: Nontransitive dice with equal means. In: Optimal Play: Mathematical Studies in Games and Gambling, pp. 293–310 (2000)
11. Gardner, M.: Paradox of nontransitive dice and elusive principle of indifference. Sci. Am. **223**, 110–111 (1970)
12. Grime, J.: The bizarre world of nontransitive dice: games for two or more players. College Math. J. **48**, 2–9 (2017)
13. John, K.: Hunter and Bruno Nachtergaele, Applied analysis. World Scientific, River Edge (2001)
14. Jan, H., Mossel, E., Ross, N., Zheng, G.: The probability of intransitivity in dice and close elections. Probab. Theory Rel. Fields **178**, 951–1009 (2020)
15. Janson, S., Łuczak, T., Rucinski, A.: Random graphs. In: Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley, New York (2000)
16. Lindeberg, J.W.: Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung. Math. Z. **15**, 211–225 (1922)
17. Meka, R., Nguyen, O., Van, V.: Anti-concentration for polynomials of independent random variables. Theory Comput. **12**, 11, 16 (2016)
18. Moon, J.W., Moser, L.: Generating oriented graphs by means of team comparisons. Pac. J. Math. **21**, 531–535 (1967)
19. Mossel, E., O'Donnell, R., Oleszkiewicz, K.: Noise stability of functions with low influences: invariance and optimality. Ann. Math. **171**, 295–341 (2010)
20. O'Donnell, R.: Analysis of Boolean Functions. Cambridge University Press, New York (2014)
21. Petrov, V.V.: Sums of Independent Random Variables, vol. 82. Springer, Berlin (2012)
22. Polymath, D.H.J.: The probability that a random triple of dice is transitive. arXiv:2211.16156
23. Schwenk, A.J.: Beware of geeks bearing grifts. Math. Horiz. **7**, 10–13 (2000)
24. Tao, T., Vu, V.H.: Additive combinatorics. In: Cambridge Studies in Advanced Mathematics, vol. 105. Cambridge University Press, Cambridge (2010)
25. Thörnblad, E.: Decomposition of tournament limits. Eur. J. Combin. **67**, 96–125 (2018)
26. Vershynin, R.: High-dimensional probability. In: Cambridge Series in Statistical and Probabilistic Mathematics, vol. 47. Cambridge University Press, Cambridge, (2018)
27. von Bahr, B., Esseen, C.-G.: Inequalities for the rth absolute moment of a sum of random variables, $1 \leq r \leq 2$. Ann. Math. Statist. **36**, 299–303 (1965)
28. Yakusheva, A.N.: Nontransitive dice with equal means and variances. Mat. Teor. Igr Prilozh. **14**, 101–120 (2022)

29. Zhao, Y.: Graph Theory and Additive Combinatorics: Exploring Structure and Randomness. Cambridge University Press, Cambridge (2023)