University Library

MINERVA ACCESS

A gateway to Melbourne's research publications

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Hazla, J;Mossel, E;Ross, N;Zheng, G

Title:
The probability of intransitivity in dice and close elections

Date:
2020-08-19

Citation:
Hazla, J., Mossel, E., Ross, N. & Zheng, G. (2020). The probability of intransitivity in dice and close elections. PROBABILITY THEORY AND RELATED FIELDS, 178 (3-4), pp.951-1009. https://doi.org/10.1007/s00440-020-00994-7.

Persistent Link:
https://hdl.handle.net/11343/252169

License:
cc-by

Probability Theory and Related Fields (2020) 178:951–1009
https://doi.org/10.1007/s00440-020-00994-7

The probability of intransitivity in dice and close elections

Jan Hązła[1] · Elchanan Mossel[2] · Nathan Ross[3] · Guangqu Zheng[4]

Received: 30 July 2019 / Revised: 1 April 2020 / Published online: 19 August 2020
© The Author(s) 2020

Abstract
We study the phenomenon of intransitivity in models of dice and voting. First, we follow a recent thread of research for n-sided dice with pairwise ordering induced by the probability, relative to $1/2$, that a throw from one die is higher than the other. We build on a recent result of Polymath showing that three dice with i.i.d. faces drawn from the uniform distribution on $\{1, \ldots, n\}$ and conditioned on the average of faces equal to $(n+1)/2$ are intransitive with asymptotic probability $1/4$. We show that if dice faces are drawn from a non-uniform continuous mean zero distribution conditioned on the average of faces equal to 0, then three dice are transitive with high probability. We also extend our results to stationary Gaussian dice, whose faces, for example, can be the fractional Brownian increments with Hurst index $H \in (0, 1)$. Second, we pose an analogous model in the context of Condorcet voting. We consider n voters who rank k alternatives independently and uniformly at random. The winner between each two alternatives is decided by a majority vote based on the preferences. We show that in this model, if all pairwise elections are close to tied, then the asymptotic probability of obtaining any tournament on the k alternatives is equal to $2^{-k(k-1)/2}$, which markedly differs from known results in the model without conditioning. We also explore the Condorcet voting model where methods other than simple majority are used for pairwise elections. We investigate some natural definitions of "close to tied" for general functions and exhibit an example where the distribution over tournaments is not uniform under those definitions.

Keywords Intransitivity · Dice · Condorcet paradox

Mathematics Subject Classification 60C05

E.M. and J.H. are partially supported by awards ONR N00014-16-1-2227, NSF CCF-1665252 and DMS-1737944. N.R. and G.Z were partially supported by ARC DP150101459. E.M. was supported by the Simons Investigator award (622132).

✉ Jan Hązła
jan.hazla@epfl.ch

Extended author information available on the last page of the article

1 Introduction

The phenomenon of intransitivity often arises when one ranks three or more alternatives. An early example is the Condorcet paradox, discovered in the eighteenth century in the context of voting. This type of intransitivity is much more general, as proved by Arrow in his social choice theorem [2]. A different fascinating aspect of intransitivity arises in the context of games of chance: The striking phenomenon of non-transitive dice. It was discovered by the statistician Brad Efron [10] and has fans such as Warren Buffet (who reportedly tried to trick Bill Gates [19]). The main motivating question of this paper is: What is the chance of observing intransitivity in natural random setups? We present some quantitative answers to this question. We introduce and discuss our results for dice and voting separately, making comparisons between the two settings where appropriate.

1.1 Intransitive dice: transitivity of non-uniform dice

For the purposes of this paper, we call an *n-sided die* (think of gambling dice) any vector $\boldsymbol{a} = (a_1, \ldots, a_n)$ of real numbers. The *face-sum* of a die $\boldsymbol{a}$ is $\sum_{i=1}^{n} a_i$. We say that die $\boldsymbol{a}$ beats die $\boldsymbol{b}$, denoted $\boldsymbol{a} \succ \boldsymbol{b}$, if a uniformly random face of $\boldsymbol{a}$ has greater value than a random face of $\boldsymbol{b}$. In other words, $\boldsymbol{a} \succ \boldsymbol{b}$ if

$$\left(\sum_{i,j=1}^{n} \mathbb{I}[a_i > b_j] - \mathbb{I}[a_i < b_j] \right) > 0.$$

We call a finite set of n-sided dice *intransitive* if the "beats" relation on the set cannot be extended to a linear order. That is, a set of dice is intransitive if it contains a subset $\boldsymbol{a}^{(1)}, \ldots, \boldsymbol{a}^{(k)}$ such that $\boldsymbol{a}^{(1)} \succ \boldsymbol{a}^{(2)} \succ \cdots \succ \boldsymbol{a}^{(k)} \succ \boldsymbol{a}^{(1)}$. A well-known example with three sides is $\boldsymbol{a} = (2, 4, 9)$, $\boldsymbol{b} = (1, 6, 8)$ and $\boldsymbol{c} = (3, 5, 7)$. One checks that $\boldsymbol{a} \succ \boldsymbol{b} \succ \boldsymbol{c} \succ \boldsymbol{a}$. If a set of dice forms a linear ordering, then we call it *transitive*. Because of ties, there can be sets that are neither transitive nor intransitive, but they occur with negligible probability in the models we study.

Recently, there has been some interest in the quantitative study of intransitive dice. The main quantity of interest is the probability that three independent dice are transitive, under different random models. In particular, as the number of faces grows, the dice can behave *transitively*, i.e., such that a triple of random dice is transitive with high probability. At the other end of the spectrum, there can be behavior that we call, borrowing the term from Kalai's paper on social choice [17], *chaotic*: in that regime, three dice are intransitive with probability[1] approaching $1/4$.

Some (mostly) experimental results were presented by Conrey, Gabbard, Grant, Liu and Morrison [7]. Among others, they conjectured that the model where n-sided dice are sampled uniformly from multisets of integers between 1 and n conditioned on the face-sum equal to $n(n+1)/2$ is chaotic. A recent collaborative Polymath project

[1] By considering paths of length two in the tournament graph on dice according to the "beats" relation, one can see that $1/4$ is the highest possible probability of intransitivity (see [31]).

[32] proved this conjecture for a related, but not identical, model where a die is a random sequence of integers between 1 and n conditioned on the face-sum equal to $n(n+1)/2$.

One may wonder what happens without the face-sum conditioning. In that case it can be seen in [31] that if the faces are only i.i.d. (with distribution depending on n), then as soon as the face-sums of dice $\boldsymbol{a}$ and $\boldsymbol{b}$ differ by significantly more than $n \log n$, the die with the higher face-sum beats the other one with high probability. In particular, three random dice with uniform faces from $\{1, \ldots, n\}$ without conditioning are transitive with high probability.

One might just as well study dice with faces drawn from a continuous probability distribution. In particular, experiments and intuition strongly suggest that the model where faces are uniform in $(-1, 1)$ and conditioned on face-sum equal zero is, as in the discrete case, chaotic.

Our first result indicates that this behavior is quite fragile. If the uniform faces are replaced with any other continuous distribution (satisfying some reasonable assumptions), then whether a die beats another is determined by the value of a real function of the faces of each die and the model becomes transitive.

Theorem 1 *Take $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ to be three independent n-sided dice with i.i.d. faces. Assume that the distribution of a single face has density* (PDF) f *and* CDF F*, mean zero and variance one. Let ε_0 denote the event that the face-sums of $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ are all zero. Additionally, assume that the distribution of a single face:*

- *Has enough* (say, six) *finite moments.*
- *Has PDF f supported on a* (possibly infinite) *closed interval* $\operatorname{supp}(f)$*. Furthermore, f is continuous on* $\operatorname{supp}(f)$.
- *Is* not *uniform on* $\left[-\sqrt{3}, \sqrt{3}\right]$.

Then:

1. *Conditional on ε_0, with probability tending to one as $n \to \infty$,*

$$\boldsymbol{a} \text{ beats } \boldsymbol{b} \text{ if and only if } \sum_{i=1}^{n} F(a_i) > \sum_{i=1}^{n} F(b_i) \ .$$

2. *As $n \to \infty$, $\mathbb{P}[\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ are transitive $\mid \varepsilon_0] \to 1$.*

To understand the differing behavior of uniform versus non-uniform dice implied by Theorem 1 and the Polymath result, we first recall that, as shown by Polymath [31], for unconditioned dice with faces uniform in $(0, 1)$, the face-sums determine if $\boldsymbol{a}$ beats $\boldsymbol{b}$ with high probability. Taking an arbitrary single-face distribution F, without conditioning on face-sums the distribution of the random variable $W = \sum_{i,j=1}^{n} \mathbb{I}[a_i > b_j]$ does not depend on F: this is because $a_i > b_j$ if and only if $F(a_i) > F(b_j)$, and since $(F(a_1), \ldots, F(a_n))$ is a die with faces uniform in $(0, 1)$; see also our Theorem 6. Therefore, considering distribution F conditioned on ε_0, for the purposes of the "beats" relation, one can just as well think of a die $(F(a_1), \ldots, F(a_n))$ conditioned on $\sum_{i=1}^{n} a_i = 0$. As long as F is not affine, one might expect that, even under ε_0, the

random variables $F(a_i)$ are distributed (almost) uniformly in $(0, 1)$ with only weak, global dependencies, suggesting that the expression

$$\operatorname{sgn}\left(\sum_{i=1}^{n} F(a_i) - F(b_i)\right)$$

still determines the winner with high probability. Note that this heuristic fails for the uniform distribution since in that case the CDF-sum is a determinstic function of the face-sum.

Applying the same reasoning in reverse, our result can be interpreted as showing that

$$\lim_{n\to\infty} \mathbb{P}\left[\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \text{ are intransitive} \mid \sum_{i=1}^{n} G(a_i) = \sum_{i=1}^{n} G(b_i) = \sum_{i=1}^{n} G(c_i) = 0\right] = 0$$

for *uniform* dice $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ for a large class of continuous, increasing, non-affine functions $G : \mathbb{R} \to \mathbb{R}$. This suggests that the intransitivity phenomenon for uniform dice is strongly linked to conditioning on the slices $\sum_{i=1}^{n} a_i = c$.

Note that the assumptions of Theorem 1 imply that the PDF f is bounded. We believe that they can be weakened in that respect: For example, it should be enough that the convolution $f^{(*k)}$ is bounded for some finite k (with the support interval $\operatorname{supp}(f)$ not necessarily closed) and that the assumption of continuity of f is replaced with piecewise continuity. We do not treat those relaxed assumptions for the sake of readability. In any case, based, among others, on experiments involving Cauchy distribution, we suspect that the first two itemized assumptions in Theorem 1 are not necessary for its statement to hold.

The main ingredient of the proof is a variance calculation that establishes that for two dice

$$\operatorname{Var}\left[\sum_{i,j=1}^{n} \mathbb{I}(a_i > b_j) - n\sum_{i=1}^{n}\big(F(a_i) - F(b_i)\big) \mid \varepsilon_0\right] = o(n^3)\,,$$

while the variance of each term of the difference is of order n^3. These two facts and an anti-concentration argument then imply Theorem 1. The variance calculation uses a CLT calculation with a rather attentive tracking of errors. This is interesting in comparison with [32], since it suggests that careful application of central limit theorems is important in establishing both transitivity and intransitivity results. We also need to establish CLT-like anti-concentration for the random variable $\sum_{i=1}^{n} F(a_i)$ conditioned on ε_0. For that, we employ a direct argument that uses conditioning on the values of pairs $a_1 + a_2, \ldots, a_{n-1} + a_n$. The proof is given in Sect. 2.

1.2 Intransitive dice: stationary Gaussian dice

In the setting of Theorem 1 with standard Gaussian $\mathcal{N}(0,1)$ faces, it can be computed that the conditioned die $\boldsymbol{a} = (a_1, \dots, a_n)$ is distributed as a joint centered Gaussian with $\mathrm{Var}[a_i] = 1 - 1/n$ and $\mathrm{Cov}[a_i, a_j] = -1/n$ for $i \neq j$. Therefore, it can be seen as a *locally stationary* Gaussian family, that is, a family where the correlation of a_i and a_j depends only on n and $i - j$ (more precisely, for our conditioning, the correlation depends solely on whether i is equal to j, i.e., δ_{ij}).

In this particular Gaussian case, one can provide another proof of the conclusion of Theorem 1 using the so-called Malliavin–Stein machinery (see [28] for a comprehensive treatment). Indeed, one can expand the indicator function $\mathbb{I}[\bullet > 0]$ based on Hermite polynomials (see (1.4)), then rewrite the random variable $W = \sum_{i,j=1}^{n} \mathbb{I}[a_i - b_j > 0]$ into an infinite sum of multiple Wiener–Itô integrals. It is then enough to apply (for example) Theorem 6.3.1 in [28] to get the following CLT:

$$\frac{1}{n^{3/2}}\Big[W - \mathbb{E}(W)\Big] \xrightarrow[n\to+\infty]{\text{law}} \mathcal{N}(0, \alpha)$$

where the limiting variance $\alpha = \frac{1}{6} - \frac{1}{2\pi}$ can be deduced from standard arguments and Newton's 1676 identity (see Remark 4). On the other hand, one can again use the Hermite expansion to compute that variance of $W - n\sum_{i=1}^{n}[F(a_i) - F(b_i)]$ is $O(n^2)$. Then the transitivity follows from this variance estimate and the above CLT. We leave the details for interested readers. Meanwhile, it is natural to investigate the (globally) stationary Gaussian case. It turns out that one can use the Breuer–Major theorem [6] to prove a version of Theorem 1 for (globally) *stationary Gaussian dice*.

Here is our setting: let $\{G_i, i \in \mathbb{N}\}$ be a centered stationary Gaussian sequence such that $\mathbb{E}[G_i G_j] = \rho(i-j)$ for some (correlation) function $\rho : \mathbb{Z} \to \mathbb{R}$. We assume that $\rho(0) = 1/2$. The main example of such a correlation function will be that of fractional Brownian increments. That is, we will consider a rich source of examples where $\rho(k) = s_H(k) := \frac{1}{2}\mathbb{E}[B_1^H(B_{|k|+1}^H - B_{|k|}^H)]$ for $k \in \mathbb{Z}$ with B^H being the *fractional Brownian motion* with Hurst parameter $H \in (0,1)$. The multiplicative constant $1/2$ is chosen only for normalization purposes and

$$s_H(k) = \frac{1}{4}\big(|k+1|^{2H} + |k-1|^{2H} - 2|k|^{2H}\big)\,; \tag{1.1}$$

one can easily check that for $H \neq 1/2$, as $|k| \to +\infty$,

$$s_H(k) \sim c_H |k|^{2H-2}\,, \tag{1.2}$$

where $c_H := H(2H-1)/2$ is uniformly bounded by $1/2$. For a brief introduction to the fractional Brownian motion, one can refer to the recent book [27].

In the following, we first present a very peculiar phenomenon arising from the fractional Brownian example as a prelude, and we postpone results concerning more general correlation functions ρ to Sect. 3.

Theorem 2 *Let $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ be i.i.d. copies of $\{G_1, \ldots, G_n\}$ with correlation function s_H for* **any** *given $H \in (0, 1)$. Then, with high probability,*

$$\boldsymbol{a} \text{ beats } \boldsymbol{b} \quad \text{if and only if} \quad \sum_{i=1}^{n} F(a_i) > \sum_{i=1}^{n} F(b_i)\,, \tag{1.3}$$

where $F(x) = \Phi(\sqrt{2}x)$ is the distribution function of $G_1 \sim N(0, 1/2)$. As a consequence, the probability that three dice $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ are transitive tends to one, as $n \to +\infty$.

Remark 1 (i) The case $H = 1/2$ corresponds to the aforementioned unconditional Gaussian dice, and by the standard integral transform, it extends to unconditional dice with i.i.d. faces sampled from a large class of distributions; see Theorem 6. As already mentioned, [31] gives an elementary proof for unconditioned uniform dice.
(ii) For $k \neq 0$, $s_H(k) > 0$ if $H \in (1/2, 1)$ while $s_H(k) < 0$ whenever $H \in (0, 1/2)$. Theorem 2 suggests that *negative correlation or positive correlation among different faces does not influence formula* (1.3), and therefore also the transitivity of $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$.

The proof of Theorem 2 makes use of the very close relation between the Hermite expansions of functions $\mathbb{I}[\bullet > 0]$ and Φ:

$$\mathbb{I}\big[\bullet > 0\big] = \frac{1}{2} + \sum_{k=0}^{\infty} d_{2k+1} H_{2k+1}, \quad \text{with } d_{2k+1} = \frac{(-1)^k}{2^k k!(2k+1)\sqrt{2\pi}}, \tag{1.4}$$

$$\Phi = \frac{1}{2} + \sum_{k=0}^{\infty} \ell_{2k+1} H_{2k+1}\,, \quad \text{with } \ell_{2k+1} = d_{2k+1} 2^{-k-\frac{1}{2}}, \tag{1.5}$$

where the above series converge in $L^2(\mathbb{R}, \exp(-x^2/2)dx)$; see Sect. 3 for more details.

1.3 Condorcet paradox: social chaos for close majority elections

The Condorcet paradox is a well-known intransitivity phenomenon in social choice theory. Consider n voters trying to decide between k alternatives. Each voter has a ranking (linear ordering) of the alternatives and we would like to aggregate the n rankings into a global one. A natural approach is as follows: given a pair of alternatives a and b, we say that a beats b if a majority of voters put a ahead of b in their rankings (we always assume n is odd to avoid dealing with ties). Aggregating these majority elections for all $K := \binom{k}{2}$ pairs of alternatives, we obtain a tournament graph on k vertices, that is, a complete graph where each edge is directed.

If there exists a Condorcet winner (i.e. the alternative that beats all others), and, in particular, if this tournament is transitive (i.e. it induces a linear ordering), we might conclude that there is a clear global winner of the election. However, in Condorcet paradox the pairwise rankings need not produce a Condorcet winner. For example, we might have three voters with rankings $a \succ b \succ c$, $b \succ c \succ a$ and $c \succ a \succ b$, respectively. Majority aggregation results in a beating b, b beating c and c beating a.

Assume a probabilistic model with n voters and k alternatives, where each voter samples one of $k!$ rankings independently and uniformly. This is called the *impartial culture* assumption and is the most common model studied in social choice (see [12] for one survey of results in related settings). Despite the example above, one might hope that under impartial culture, the paradox is unlikely to arise for a large number of voters. However, it was one of the earliest results in social choice theory [11,13] that it is not so: in particular, letting $P_{\text{Cond}}(k, n)$ to be the probability of Condorcet winner for n voters and k alternatives, and $P_{\text{Cond}}(k) := \lim_{n\to\infty} P_{\text{Cond}}(k, n)$, we have

$$P_{\text{Cond}}(3) = \frac{3}{2\pi}\arccos(-1/3) \leq 91.2\% \ . \tag{1.6}$$

For $k \geq 4$ there is no simple expression, but the numerical values up to $k = 50$ were computed by Niemi and Weisberg [26]; for example, $P_{\text{Cond}}(10) \approx 51.1\%$ and $P_{\text{Cond}}(27) \approx 25.5\%$, and the asymptotic behavior is given by May [21] as

$$P_{\text{Cond}}(k) = \frac{\sqrt{8\pi \log k}}{k}\big(1 + O(1/\log k)\big) \ , \tag{1.7}$$

in particular $\lim_{k\to\infty} P_{\text{Cond}}(k) = 0$. If one is interested in the probability of a completely transitive outcome, the best asymptotic estimate known [22] is $\exp(-\Theta(k^{5/3}))$.

Given the dice models studied in [7] and [32], it seems reasonable to study the probability of Condorcet paradox under impartial culture, conditioned on all pairwise elections being close to tied. The conditioning on elections being almost tied seems natural also given the abundance of real life elections that are close to tied.

To define the model more precisely, for each pair of alternatives $\{a, b\}$, define the random variable $S^{(ab)}$ to be the number of voters that prefer a to b, minus the number of voters preferring b to a. In other words, the sign of $S^{(ab)}$ determines the alternative that wins the pairwise election. Let $Y^{(ab)} := \text{sgn}(S^{(ab)})$ and Y be the random tuple encoding the K pairwise winners via the $Y^{(ab)}$, having K entries with values in $\{-1, 1\}$. Furthermore, for $d \geq 1$, let ε_d be the event that $\left|S^{(ab)}\right| \leq d$ for every pair $\{a, b\}$. We think of the event ε_d as "the elections are d-close", with $d = 1$ corresponding to almost perfectly tied elections.

Our main result for voting uses a multidimensional local limit theorem to show that the probability of Condorcet winner for almost tied elections goes to zero much faster than in (1.7). Actually, we prove the following stronger result.

Theorem 3 *Let n be odd, $d \geq 1$ and $y \in \{-1, 1\}^K$. Then,*

$$\left|\mathbb{P}\left[Y = y \mid \varepsilon_d\right] - \frac{1}{2^K}\right| \leq \alpha_k \frac{d^2}{n} + o_k(1) \ , \tag{1.8}$$

where $\alpha_k > 0$ depends only on k and $o_k(1)$ denotes a function that depends only on k (but not on d or y) and goes to zero, as n goes to infinity.

In particular,

$$\left|\mathbb{P}\left[Y \text{ is transitive } \mid \varepsilon_d\right] - \frac{k!}{2^K}\right| \leq \beta_k \frac{d^2}{n} + o_k(1) \tag{1.9}$$

and

$$\left|\mathbb{P}\left[Y \text{ has Condorcet winner} \mid \varepsilon_d\right] - \frac{k}{2^{k-1}}\right| \leq \gamma_k \frac{d^2}{n} + o_k(1) \tag{1.10}$$

for some $\beta_k, \gamma_k > 0$.

One interpretation of this result is that the probability of Condorcet paradox, which is already substantial without conditioning, increases to reach the fully chaotic behavior for elections that are almost three-way ties. The event ε_d for $d = o(\sqrt{n})$ has subconstant probability, but on the other hand such "close" elections seem to be a natural case to study (and one might argue that in practice they arise more often than the model suggests). Furthermore, some other interesting phenomena in social choice can be shown to arise only with polynomially small probability, see, e.g. the quantitative Gibbard–Satterthwaite theorem [9,14,25].

Comparing Theorem 3 to intransitivity of random uniform dice conditioned on their face-sums, first note that for almost tied elections and $k = 3$, the asymptotic probability of Condorcet winner computed from (1.10) is $3/4$, which is equal to the probability of transitivity for dice. On the other hand, there is a difference in the transition between the transitive and chaotic regimes. Assuming dice with faces uniform in $(-1, 1)$, the model is chaotic when conditioned on face-sums equal to zero, but, as shown by Polymath [31], it becomes transitive as soon as we condition on face-sums of absolute value at most d for $d = \omega(\log n)$. However, the voting outcomes behave chaotically for d-close elections for any $d = o(\sqrt{n})$ and transition into the "intermediate", rather than transitive, regime given by (1.6). Furthermore, (1.8) means that the tournament on k alternatives determined by Y is asymptotically random. [7] conjectured that k random dice also form a random tournament, however [32] report experimental evidence against this conjecture.

We also note that the proof of Theorem 3 can be modified such that its statement holds even when conditioning on only $K-1$ out of K pairwise elections being d-close.

The above-mentioned work by Kalai [17] calls the situation when Y is a random tournament *social chaos*. He considers impartial culture model (without conditioning) and an arbitrary monotone odd function $f : \{-1, 1\}^n \to \{-1, 1\}$ for pairwise elections (the setting we considered so far corresponds to $f = \mathrm{Maj}_n$). Under these assumptions, he proves that social chaos is equivalent to the asymptotic probability of Condorcet winner for three alternatives being equal to $3/4$. [17] contains another equivalent condition for social chaos, stated in terms of noise sensitivity of function f for only two alternatives. It is interesting to compare it with the reduction from three to two dice in Lemma 2.1 of [32].

1.4 Condorcet paradox: generalizing close elections—a case study

It would be interesting to extend Theorem 3 to other natural pairwise comparison functions such as weighted majorities and recursive majorities, similar to the electoral college in the USA. However, in order to formulate such a result, it is first necessary to define d-close elections for an arbitrary function. The results of this section deal with the question if such a definition exists. Somewhat surprisingly, we show that natural definitions of close elections do not lead to a chaotic outcome when ranking three alternatives. We do so by presenting a simple example, for which two of the most natural definitions do not result in chaotic outcome.

For this we consider the following function. Let us assume that there are three candidates a, b, c and a number of voters n that is divisible by three, letting $m := n/3$. We take $f : \{-1, 1\}^n \to \{-1, 1\}$ to be

$$f(x_1, \ldots, x_n) := \operatorname{sgn}\left(\sum_{i=1}^{m} \operatorname{sgn}\left(x_{3i-2} + x_{3i-1} + x_{3i}\right)\right).$$

In words, f is a two-level majority: majority of votes of m triplets, where the vote of each triplet is decided by majority.

The function f possesses many pleasant properties: it is odd, transitive symmetric[2] and is a polynomial threshold function of degree three. We would like to devise a natural notion of d-close elections according to f. In light of Theorem 3 it might be argued that the "right" notion of closeness should result in the chaotic outcome, same as for majority. We show that for two natural definition of closeness, *this is not the case*.

To start with, let $w_i := x_{3i-2} + x_{3i-1} + x_{3i}$. In the following we will sometimes treat f as a function of $\mathbf{w} := (w_1, \ldots, w_m)$, i.e., $f : \{\pm 1, \pm 3\}^m \to \{\pm 1\}$, with the distribution of $\mathbf{w}$ induced by the distribution of $\mathbf{x}$, i.e., $w_i = \pm 3$ and $w_i = \pm 1$ with probabilities $1/8$ and $3/8$, respectively. A CLT argument as in Theorem 3 implies chaotic behavior of f if we define "d-close" as "$\left|\sum_{i=1}^{m} \operatorname{sgn}\left(w_i^{(kk')}\right)\right| \le d$" for every pair of candidates (kk'). However, this is not very satisfactory for at least two reasons. First, it does not seem to extend to other functions that do not have such an "obvious" summation built into them. Second, it does not accord well with our intuition of closeness. This second problem becomes more apparent considering analogous condition for another two-level majority, with $\sqrt{n}$ groups of $\sqrt{n}$ voters each. In this case of "electoral college" an election that was close in every "state" in favor of a single candidate would not be considered close overall.

Another idea is to define "d-close" the same way as in Theorem 3, that is as "$\left|\sum_{i=1}^{n} x_i^{(kk')}\right| \le d$". Clearly, this is not a good closeness measure for an arbitrary comparison method (e.g., weighted majority with large differences between weights), but one could argue that it is relevant at least for transitive symmetric functions. Using another CLT argument, we find that for this definition of closeness, the behavior

[2] A voting function $f : \{-1, 1\}^n \to \{-1, 1\}$ is transitive symmetric if for every $i, j \in [n]$ there exists a permutation $\sigma : [n] \to [n]$ such that $\sigma(i) = j$ and $f \circ \sigma = f$, where $(f \circ \sigma)(x_1, \ldots, x_n) = f(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$. Informally, every two voters play the same role.

of $o(\sqrt{n})$-close elections under f is not chaotic: the asymptotic Condorcet paradox probability is slightly less than 25%. Note that for three candidates, the Condorcet paradox occurs if and only if $f(\mathbf{x}^{(ab)}) = f(\mathbf{x}^{(bc)}) = f(\mathbf{x}^{(ca)})$.

Theorem 4 *Under the notation above and the event ε_d as defined in Sect. 1.3, for $d = \sqrt{n}/\log n$,*

$$\lim_{n\to\infty} \mathbb{P}\left[f(\boldsymbol{x}^{(ab)}) = f(\boldsymbol{x}^{(bc)}) = f(\boldsymbol{x}^{(ca)}) \mid \varepsilon_d\right] = \alpha^* ,$$

where $\alpha^* \approx 23.2\%$ is an absolute constant.

For comparison, without conditioning the Condorcet paradox probability is $\approx 12.5\%$ when the elections are according to f and $\approx 8.8\%$ according to majority.

The idea for the proof of Theorem 4 is to use multivariate Berry–Esseen theorem for random variables

$$\left(A^{(kk')}, B^{(kk')}\right)_{(kk')} := \left(\sum_{i=1}^{n} x_i^{(kk')}, \sum_{i=1}^{m} \operatorname{sgn}\left(w_i^{(kk')}\right)\right)_{(kk')}, \quad kk' \in \{ab, bc, ca\}.$$

We are looking at sign patterns of $B^{(kk')}$ conditioned on small absolute values of $A^{(kk')}$. $A^{(kk')}$ and $B^{(kk')}$ are not perfectly correlated and it turns out that part of (negative) correlations between $B^{(ab)}$, $B^{(bc)}$ and $B^{(ca)}$ is not attributable to correlations between $A^{(ab)}$, $A^{(bc)}$ and $A^{(ca)}$. Hence, even after conditioning on small $A^{(kk')}$ there remains a small constant correlation between $B^{(kk')}$, which prevents completely chaotic behavior.

Another promising definition of closeness involves the noise operator T_ρ from the analysis of Boolean functions (see e.g., [29] for more details). Let $\rho \in [-1, 1]$ and $\mathbf{x} \in \{-1, 1\}^n$. Define a probability distribution $N_\rho(\mathbf{x})$ over $\{-1, 1\}^n$ such that $y_1, \ldots, y_n$ are sampled independently with $y_i = -x_i$ with probability $\varepsilon := \frac{1-\rho}{2}$ and $y_i = x_i$ otherwise. Note that $\mathbb{E}[x_i y_i] = \rho$, hence we say that a pair $(\mathbf{x}, \mathbf{y})$ sampled as uniform $\mathbf{x}$ and then $\mathbf{y}$ according to $N_\rho(\mathbf{x})$ is *ρ-correlated*. The *noise operator* T_ρ is defined as

$$T_\rho f(\mathbf{x}) := \mathbb{E}_{\mathbf{y}\sim N_\rho(\mathbf{x})}\left[f(\mathbf{y})\right] .$$

For $\rho \in (0, 1)$ one can think of $N_\rho(\mathbf{x})$ as a distribution over $\{-1, 1\}^n$ with probabilities that are decreasing in the Hamming distance from $\mathbf{x}$. Furthermore, for f being majority and $d = o(\sqrt{n})$ the condition $\left|\sum_{i=1}^n x_i\right| \le d$ is asymptotically equivalent to $\left|T_\rho \mathrm{Maj}(\mathbf{x})\right| \le C_\rho d/\sqrt{n}$. This suggests that it may be fruitful to define "d-close" as "$|T_\rho f(\mathbf{x}^{(kk')})| \le d/\sqrt{n}$". The idea becomes even more appealing when considering a Fourier-analytic Condorcet formula discovered by Kalai [16]. He showed that for an odd function $g\colon \{-1, 1\}^n \to \{-1, 1\}$, the probability of Condorcet paradox without

conditioning is equal to

$$\mathbb{P}\left[g(\mathbf{x}^{(ab)}) = g(\mathbf{x}^{(bc)}) = g(\mathbf{x}^{(ca)})\right] = \frac{1}{4}\left(1 - 3\,\mathbb{E}_{\mathbf{x},\mathbf{y}}\left[g(\mathbf{x})g(\mathbf{y})\right]\right)$$
$$= \frac{1}{4}\left(1 - 3\,\mathbb{E}_{\mathbf{x}}\left[g(\mathbf{x})T_{1/3}g(\mathbf{x})\right]\right), \qquad (1.11)$$

where $(\mathbf{x}, \mathbf{y})$ are $1/3$-correlated.

Another feature of the T_ρ operator is that for noise sensitive functions (which [17] proved to be exactly those that result in chaotic elections without conditioning) the value $|T_\rho f(\mathbf{x})|$ is $o(1)$ with high probability over $\mathbf{x}$. If we decide to use $|T_\rho f(\mathbf{x})|$ as a measure of closeness, then this fact can be given the following (though by no means the only possible) interpretation: elections held according to a noise sensitive function are almost always close.

Recall our "majority of triplets" function f and define the event $\mathcal{F}_{\rho,d}$ as

$$\mathcal{F}_{\rho,d} :\equiv \quad \max\left(\left|T_\rho f(\mathbf{x}^{(ab)})\right|, \left|T_\rho f(\mathbf{x}^{(bc)})\right|, \left|T_\rho f(\mathbf{x}^{(ca)})\right|\right) \le \frac{d}{\sqrt{m}}\,.$$

At first sight, (1.11) suggests that the event $\mathcal{F}_{\rho,d}$, with $\rho = 1/3$ and $d = o(\sqrt{m})$, should cause the expectation term in (1.11) to vanish and the probability of Condorcet paradox to approach $1/4$. Surprisingly, this is not the case for f:

Theorem 5 *Fix $\rho \in (0, 1)$ and take $d := \sqrt{m}/\log m$. Then,*

$$\lim_{n\to\infty} \mathbb{P}\left[f(\boldsymbol{x}^{(ab)}) = f(\boldsymbol{x}^{(bc)}) = f(\boldsymbol{x}^{(ca)}) \mid \mathcal{F}_{\rho,d}\right] = \alpha(\rho)\,,$$

where $\alpha(\rho) \in [0.17, \alpha^*]$ with α^* the constant from Theorem 4 and $\alpha(\rho) \to \alpha^*$ as $\rho \to 0^+$.

The proof of Theorem 5 is a variation on the proof of Theorem 4. For $\mathbf{w} \in \{\pm 3, \pm 1\}^m$ and $b \in \{\pm 3, \pm 1\}$, we let $W_b(\mathbf{w}) := |\{i \in [m] : w_i = b\}|$ and $V_b(\mathbf{w}) := W_b(\mathbf{w}) - \mathbb{E}_{\mathbf{w'}}[W_b(\mathbf{w'})]$. Then, we observe that, just as for majority the value of $T_\rho \mathrm{Maj}(\mathbf{x})$ is proportional to the number of ones in $\mathbf{x}$ minus $n/2$, also for f the value of $T_\rho f(\mathbf{w})$ is proportional to a certain linear combination of $V_b(\mathbf{w})$. This allows us to proceed with an identical argument as in Theorem 4 with appropriately redefined random variables $A^{(kk')}$.

Some more recent results show that, without conditioning, majority in fact maximizes the probability of Condorcet winner among "low-influence functions" (see [24] for three voters and [15,22] for general case). This contrasts with Theorems 4 and 5 for different definitions of close elections.

1.5 Arrow's theorem for dice

To further consider the parallels between dice and social choice, we also ask if there is a dice analogue of Arrow's theorem (and its quantitative version). We obtain a rather

generic statement that does not use any properties of dice and a quantitative version which is a restatement of a result on tournaments by Fox and Sudakov [8].

Organization of the paper The proofs of our main theorems are located in Sects. 2 (Theorem 1), 3 (Theorem 2), 4 (Theorem 3) and 5 (Theorems 4 and 5). Section 6 contains the discussion of Arrow's theorem for dice. The sections are mostly self-contained and can be read in any order.

2 Transitivity of non-uniform dice

In this section we are going to prove Theorem 1. Let us start with some notation. For the sake of readability, in this section we drop the bold typesetting for dice vectors. We let

$$W_{ij}^{(kk')} := \mathbb{I}(k_i > k'_j)$$

for $k, k' \in \{a, b, c\}$ and

$$W^{(kk')} = \sum_{i,j=1}^{n} W_{ij}^{(kk')}.$$

We also let $V^{(kk')} := \sum_{i=1}^{n} F(k_i) - F(k'_i)$. An important value that we will use is

$$A := \mathbb{E}[a_1 F(a_1)]\,. \tag{2.1}$$

The constant A is significant because it distinguishes the uniform distribution: by Cauchy–Schwarz we have

$$A^2 = \mathbb{E}[a_1 F(a_1)]^2 = \mathbb{E}[a_1 (F(a_1) - 1/2)]^2 \le \mathrm{Var}[a_1] \cdot \mathrm{Var}[F(a_1)] = \frac{1}{12}$$

(note that $F(a_1)$ is uniform in $(0, 1)$, so $\mathbb{E}[F(a_1)] = 1/2$ and $\mathrm{Var}[F(a_1)] = 1/12$). On the other hand, since a_1 and $F(a_1)$ are linearly dependent if and only if distribution of a_1 is uniform on $(-\sqrt{3}, \sqrt{3})$, the equality $A^2 = 1/12$ is achieved exactly for the uniform distribution. In the non-uniform case, this leads to a key cancellation leading to (2.2) below.

Since for a non-uniform distribution clearly we have

$$\mathbb{P}\left[\sum_{i=1}^{n} F(k_i) = \sum_{i=1}^{n} F(k'_i) \mid \varepsilon_0\right] = 0$$

(see also the proof of Proposition 2), the second statement of Theorem 1 follows from the first. What needs to be done can be summed up in two propositions. In

the following proof we assume conditioning on ε_0 and drop it from the notation for readability. We also note that constants hidden in $O(\cdot)$, $o(\cdot)$, etc., are allowed to depend on the distribution F.

Proposition 1

$$Var\left[W^{(ab)} - nV^{(ab)}\right] = o(n^3)\ . \tag{2.2}$$

Proposition 2 *For every $C \in \mathbb{R}$ and $\varepsilon > 0$,*

$$\mathbb{P}\left[\frac{V^{(ab)}}{\sqrt{n}} \in [C-\varepsilon, C+\varepsilon]\right] = O(\varepsilon) + O\left(\frac{1}{\sqrt{n}}\right)\ , \tag{2.3}$$

where the $O(\cdot)$ constants do not depend on C or ε.

We note that during the proof of Proposition 1 we establish $\mathrm{Var}[W^{(ab)}]$, $\mathrm{Var}[nV^{(ab)}] \geq \Omega(n^3)$, so indeed Proposition 1 is saying that these two random variables are closely correlated.

Proof (*Theorem 1 follows from the propositions*) Let $\overline{W}^{(kk')} := W^{(kk')} - \mathbb{E}[W^{(kk')}] = W^{(kk')} - n^2/2$. It is enough to prove that

$$\mathbb{P}\left[\mathrm{sgn}\left(V^{(ab)}\right) \neq \mathrm{sgn}\left(\overline{W}^{(ab)}\right)\right] = o(1)\ .$$

For any $\delta > 0$, note that $\mathrm{sgn}\left(V^{(ab)}\right) \neq \mathrm{sgn}\left(\overline{W}^{(ab)}\right)$ implies that

$$\text{either } \left|nV^{(ab)} - \overline{W}^{(ab)}\right| > \delta \text{ or } \left|nV^{(ab)}\right| \leq \delta.$$

Furthermore, by Chebyshev's inequality and (2.2),

$$\mathbb{P}\left[\left|\overline{W}^{(ab)} - nV^{(ab)}\right| > \delta\right] < \frac{o(n^3)}{\delta^2}\ .$$

Taking appropriate $\delta := o(n^{3/2})$, we finally compute

$$\begin{aligned}
&\mathbb{P}\left[\mathrm{sgn}\left(V^{(ab)}\right) \neq \mathrm{sgn}\left(\overline{W}^{(ab)}\right)\right] \\
&\quad\leq \mathbb{P}\left[\left|nV^{(ab)} - \overline{W}^{(ab)}\right| > \delta\right] + \mathbb{P}\left[\left|nV^{(ab)}\right| \leq \delta\right] \\
&\quad= o(1) + O\left(\frac{\delta}{n^{3/2}}\right) = o(1)\ ,
\end{aligned}$$

where we used (2.3) in the last line. □

Remark 2 It is also true that with high probability a beats b if and only if $\sum_{i=1}^{n} F_n(a_i) > \sum_{i=1}^{n} F_n(b_i)$, where F_n is the CDF of the *conditional marginal* of a_1 (or any a_i) conditioned on ε_0, rather than the unconditional marginal F as in Theorem 1. (Some numerical experiments suggest that F_n is a better predictor of the "strength" of a die than F.) To see why this is true, if $V'^{(ab)} := \sum_{i=1}^{n} F_n(a_i) - F_n(b_i)$, then similar calculations to those in the proof of Proposition 1 yield

$$\operatorname{Var}\left[V'^{(ab)} - V^{(ab)}\right] = o(n) ,$$

and using this in the bound

$$\begin{aligned}&\mathbb{P}\left[\operatorname{sgn}\left(V'^{(ab)}\right) \neq \operatorname{sgn}\left(\overline{W}^{(ab)}\right)\right]\\ &\leq \mathbb{P}\left[\left|nV^{(ab)} - \overline{W}^{(ab)}\right| > \delta\right] + \mathbb{P}\left[\left|nV^{(ab)} - nV'^{(ab)}\right| > \delta\right] + \mathbb{P}\left[\left|nV^{(ab)}\right| \leq \delta\right],\end{aligned}$$

the result follows similar to above.

We proceed to prove the propositions, starting with the shorter proof of Proposition 2. In both proofs we do *not* assume conditioning on ε_0 by default.

2.1 Proof of Proposition 2

For simplicity we will assume that $n = 2m$. The idea of the proof is as follows: First, by independence, it is enough to establish anti-concentration for the single-die random variable $\sum_{i=1}^{n} F(a_i)$. Since the single-face distribution is not uniform, there must exist two points $x^*, y^* \in \operatorname{supp}(f)$ such that

$$F(x^*) + F(y^*) \neq 2F(z^*) , \tag{2.4}$$

where $z^* := \frac{x^*+y^*}{2}$. Consider random variables $d_1, \ldots, d_m$ given by

$$d_i := a_{2i-1} + a_{2i} . \tag{2.5}$$

By a concentration argument, with high probability, for a constant fraction of coordinates $i \in \{1, \ldots, m\}$, it must be that $d_i \approx 2z^*$. Furthermore, after conditioning on $d_1, \ldots, d_m$, for each such coordinate it must be that for $d_i \approx 2z^*$, both

$$\begin{aligned}&a_{2i-1} \approx x^*, a_{2i} \approx y^*,\\ &a_{2i-1}, a_{2i} \approx z^*,\end{aligned} \tag{2.6}$$

are possible with constant probability. But (2.4) and (2.6) imply that, even conditioned on $d_1, \ldots, d_m$, the variance of $\sum_{i=1}^{n} F(a_i)$ is at least $\Omega(n)$, and that allows us to apply Berry–Esseen theorem to establish a (conditional) CLT and anti-concentration. Below we present this argument in more detail, starting with an auxiliary concentration lemma.

Lemma 1 *Let $x \in \operatorname{supp}(f)$ and $\delta > 0$. There exist constants $\alpha := \alpha(f, \delta) > 0$, $\beta := \beta(f, \delta) > 0$ such that*

$$\mathbb{P}\big[\,|\{i \in [n] : x - \delta \le a_i \le x + \delta\}| < \alpha n \mid \varepsilon_0\big] \le O\left(\exp\left(-\beta n\right)\right) . \tag{2.7}$$

Proof We will think of sampling $a_1, \ldots, a_n$ conditioned on ε_0 as an experiment on $n - k$-dimensional space for some $k \in \mathbb{N}$, where the density of $(a_1, \ldots, a_{n-k})$ is proportional to $\prod_{i=1}^{n-k} f(a_i) \cdot f^{(*k)}(-a_0)$, with $a_0 := \sum_{i=1}^{n-k} a_i$ and $f^{(*k)}$ being the k-fold convolution of the PDF f.

Take $\varepsilon > 0$ and consider a set

$$I_{k,\varepsilon} := \left\{x \in \mathbb{R} : f^{(*k)}(x) > \varepsilon\right\} .$$

Since f is continuous and its support is an interval that necessarily contains zero, it must be that for every $L > 0$ there exist k large enough and ε small enough such that we have the inclusion

$$[-L, L] \subseteq I_{k,\varepsilon} .$$

We take such large enough L (as soon specified) and fix k and ε accordingly. Consider the i.i.d. choice of $a_1, \ldots, a_{n-k}$. By the Berry–Esseen theorem,

$$\begin{aligned}\mathbb{P}_{a_1,\ldots,a_{n-k}}\left[-L \le -a_0 \le L\right] &= \mathbb{P}\left[\frac{-L}{\sqrt{n-k}} \le g \le \frac{L}{\sqrt{n-k}}\right] + O\left(\frac{1}{\sqrt{n}}\right) \\ &= \Omega\left(\frac{1}{\sqrt{n}}\right) ,\end{aligned} \tag{2.8}$$

where g is a standard Gaussian random variable, and the last equality uses that L can be chosen large enough to overcome the (potentially negative) error in the normal approximation.

Let $\mathcal{F}$ be the event from (2.7), the probability of which we are bounding and define another event $\mathcal{F}'$ as

$$\mathcal{F}' :\equiv |\{i \in [n-k] : x - \delta \le a_i \le x + \delta\}| < \alpha n .$$

Taking M to be an upper bound on $f^{(*k)}(y)$ for $y \in \mathbb{R}$ and setting $\alpha := \mathbb{P}(x - \delta \le a_1 \le x + \delta)/2$, we compute

$$\begin{aligned}\mathbb{P}\left[\mathcal{F} \mid \varepsilon_0\right] &\le \mathbb{P}\left[\mathcal{F}' \mid \varepsilon_0\right] \\ &= \frac{\int \cdots \int f(a_1) \cdots f(a_{n-k}) \cdot f^{(*k)}(-a_0) \cdot \mathbb{I}[\mathcal{F}']\, \mathrm{d}a_1 \cdots \mathrm{d}a_{n-k}}{\int \cdots \int f(a_1) \cdots f(a_{n-k}) \cdot f^{(*k)}(-a_0)\, \mathrm{d}a_1 \cdots \mathrm{d}a_{n-k}} \\ &\le \frac{M \cdot \mathbb{P}_{a_1,\ldots,a_{n-k}}[\mathcal{F}']}{\varepsilon \cdot \mathbb{P}_{a_1,\ldots,a_{n-k}}[-L \le -a_0 \le L]} \\ &\le O\left(\sqrt{n}\right) \cdot \exp\left(-\beta n\right) \le O\left(\exp(-\beta' n)\right) ,\end{aligned}$$

where in the last line we used a standard Chernoff bound, since the random variable

$$|\{i \in [n-k] : x - \delta \leq a_i \leq x + \delta\}|$$

can be written as a sum of $n-k$ i.i.d. Bernoulli random variables with mean $2\alpha > 0$. □

We continue with the proof of Proposition 2, following the plan from the beginning of the section. For now, we will focus only on one half of the expression $V^{(ab)}$, namely the sum $\sum_{i=1}^{n} F(a_i)$.

Recall that by (2.4) we have $x^*, y^*, z^* = (x^* + y^*)/2$ such that

$$\gamma := |F(x^*) + F(y^*) - 2F(z^*)| > 0.$$

Furthermore, since F is continuous, we can assume that both x^* and y^* lie in the interior of the support of f. Take small $\delta > 0$ such that

$$[x^* - \delta, x^* + \delta], [y^* - \delta, y^* + \delta], [z^* - \delta, z^* + \delta] \subseteq \mathrm{supp}(f)$$

and, at the same time,

$$\begin{aligned} |w - x^*| \leq 2\delta &\implies |F(w) - F(x^*)| \leq \gamma/10, \\ |w - y^*| \leq 2\delta &\implies |F(w) - F(y^*)| \leq \gamma/10, \\ |w - z^*| \leq 2\delta &\implies |F(w) - F(z^*)| \leq \gamma/10. \end{aligned}$$

Recall the random variables $d_1, \ldots, d_m$ that we defined in (2.5). Note that the distribution of $d_1/\sqrt{2} = (a_1 + a_2)/\sqrt{2}$ satisfies the assumptions of Theorem 1. Therefore, we can apply Lemma 1 to $d_1, \ldots, d_m$, $x = 2z^* \in \mathrm{supp}(f^{(*2)})$ and δ to obtain that except with probability $\exp(-\Omega(n))$, we have that, conditioned on ε_0,

$$\left|\left\{i \in [m] : 2z^* - \delta \leq d_i \leq 2z^* + \delta\right\}\right| \geq \Omega(n)\ . \tag{2.9}$$

Observe that the distribution $a_1, \ldots, a_n$ conditioned on ε_0 can be obtained by first sampling $d_1, \ldots, d_m$ conditioned on $\sum_{i=1}^{m} d_i = 0$ and then sampling a_{2i-1} and a_{2i} conditioned on $a_{2i-1} + a_{2i} = d_i$ independently for each $i \in [m]$.

Fix a choice of $d_1, \ldots, d_m$ satisfying (2.9). We will call $i \in [m]$ that fulfills the condition from (2.9) *good*. We will now show that any such good i assumes values from (2.6) with constant probability. To that end, let us assume without loss of generality that d_1 is good and consider $d \in [2z^* - \delta, 2z^* + \delta]$. We compute (where $o(1)$ is a

function that uniformly goes to zero as δ goes to zero)

$$\begin{aligned}\mathbb{P}\left[x^*-\delta\le a_1\le x^*+\delta \mid a_1+a_2=d\right] &= \frac{\int_{x^*-\delta}^{x^*+\delta} f(x)f(d-x)\,\mathrm{d}x}{\int_{\mathbb{R}} f(x)f(d-x)\,\mathrm{d}x}\\ &\ge \frac{\int_{x^*-\delta}^{x^*+\delta}(f(x^*)+o(1))(f(y^*)+o(1))\,\mathrm{d}x}{\max_{d\in[2z^*-\delta,2z^*+\delta]} f^{(2)}(d)}\\ &\ge c\cdot\delta f(x^*)f(y^*)\ge c'>0\,, \end{aligned} \tag{2.10}$$

where c' is a positive constant achieved for small enough δ. A similar argument gives

$$\mathbb{P}\left[z^*-\delta\le a_1\le z^*+\delta \mid a_1+a_2=d\right]\ge c'>0. \tag{2.11}$$

Observe that $a_1\in[x^*-\delta,x^*+\delta]$ implies $|F(a_1)-F(x^*)|\le\gamma/10$, $a_2\in[y^*-2\delta,y^*+2\delta]$, $|F(a_2)-F(y^*)|\le\gamma/10$ and finally

$$|F(a_1)+F(a_2)-F(x^*)-F(y^*)|\le\gamma/5,$$

giving the overall conclusion

$$\mathbb{P}\Big[F(a_1)+F(a_2)\le F(x^*)+F^(y^*)+\gamma/5 \mid a_1+a_2=d\Big]\ge c'. \tag{2.12}$$

Similarly, $a_1\in[z^*-\delta,z^*+\delta]$ implies $a_2\in[z^*-2\delta,z^*+2\delta]$ and consequently

$$\big|F(a_1)+F(a_2)-2F(z^*)\big|\le\gamma/5,$$

in particular

$$F(a_1)+F(a_2)\ge 2F(z^*)-\gamma/5\ge F(x^*)+F(y^*)+\gamma/5+\gamma/2$$

and

$$\mathbb{P}\Big[F(a_1)+F(a_2)\ge F(x^*)+F(y^*)+\gamma/5+\gamma/2 \mid a_1+a_2=d\Big]\ge c'. \tag{2.13}$$

Bounds in (2.12) and (2.13) together imply that for any good i we can uniformly lower bound the conditional variance

$$\mathrm{Var}\left[F(a_{2i-1})+F(a_{2i}) \mid a_{2i-1}+a_{2i}=d_i\right]\ge\Omega(\gamma^2)\ge\Omega(1).$$

Since after conditioning on $d_1,\dots,d_m$ satisfying (2.9), the random variables $F(a_{2i-1})+F(a_{2i})$ are bounded and independent with total variance $\Omega(m)$, we can apply Berry–Esseen theorem and anti-concentration properties of a standard Gaussian

to obtain

$$\begin{aligned}\mathbb{P}\left[C-\varepsilon \le \sum_{i=1}^{n} \frac{F(a_i)}{\sqrt{n}} \le C+\varepsilon \;\middle|\; d_1,\dots,d_m\right] \\ = \mathbb{P}\left[C-\varepsilon \le \sum_{i=1}^{m} \frac{F(a_{2i-1})+F(a_{2i})}{\sqrt{2m}} \le C+\varepsilon \;\middle|\; d_1,\dots,d_m\right] \\ \le O(\varepsilon)+O\left(\frac{1}{\sqrt{n}}\right).\end{aligned}$$

Actually, since the sums $\sum_{i=1}^{n} F(a_i)$ and $\sum_{i=1}^{n} F(b_i)$ are independent even after conditioning on ε_0, we also get

$$\mathbb{P}\left[C-\varepsilon \le \frac{V^{(ab)}}{\sqrt{n}} \le C+\varepsilon \;\middle|\; d_1,\dots,d_m,d'_1,\dots,d'_m\right] \le O(\varepsilon)+O\left(\frac{1}{\sqrt{n}}\right).$$

where $d'_i = b_{2i-1}+b_{2i}$ and $d'_1,\dots,d'_m$ satisfy condition (2.9). Finally, we get (2.3) by averaging over $d_1,\dots,d_m,d'_1,\dots,d'_m$ and absorbing exponentially small terms coming from the choices that do not satisfy (2.9). □

Remark 3 One could also prove a variant of Proposition 2 by a two-dimensional local CLT argument. For example, Theorem 19.1 in [5] could be applied to show that $V^{(ab)}/\sqrt{n}$ conditioned on ε_0 converges in law to a Gaussian. However, to apply [5] it needs to be shown that there exists a finite k such that the joint distribution of

$$\left(\sum_{i=1}^{k} a_i, \sum_{i=1}^{k} F(a_i)\right)$$

has bounded density. Note that since $F(a_i)$ is a deterministic function of a_i, for $k=1$ the density does not exist. In some cases it is not difficult to show that a small $k>1$ is enough. For example, for a shifted exponential distribution with the PDF

$$f(x) = \exp(-x-1)$$

for $x \in [-1,+\infty)$ we can see that $(a_1+a_2, F(a_1)+F(a_2))$ has bounded density since the equation system

$$\begin{aligned}a_1+a_2 &= a\\ F(a_1)+F(a_2) &= a'\end{aligned}$$

has at most one solution for every pair (a,a'). On the other hand, a distribution with support $[-2,2]$ that is (up to normalization) uniform on $[-2,-1]\cup[1,2]$ and Gaussian on $(-1,1)$ does not have bounded density for any finite k.

2.2 Proof of Proposition 1

We prove Proposition 1 by a somewhat tedious computation. Recall that in this proof we do *not* assume conditioning on ε_0 by default. Also, for $k \in \{a, b, c\}$, we will denote by ε_k the single die event $\sum_{i=1}^{n} k_i = 0$.

The variance we are looking at can be broken down as

$$\operatorname{Var}\left[W - n\sum_{i=1}^{n} F(a_i) - F(b_i) \mid \varepsilon_0\right] = n^2 \operatorname{Var}\left[\sum_{i=1}^{n} F(a_i) - F(b_i) \mid \varepsilon_0\right] + \operatorname{Var}[W \mid \varepsilon_0] - 2n \sum_{i,j,k=1}^{n} \mathbb{E}\left[\mathbb{I}(a_i > b_j) \cdot (F(a_k) - F(b_k)) \mid \varepsilon_0\right]. \tag{2.14}$$

The idea is to subdivide each of the three terms above into yet smaller pieces, each of which can be written down as a certain probability involving (conditioned and unconditioned) die faces. For example,

$$\mathbb{E}\left[\mathbb{I}(a_1 > b_1) F(a_2) \mid \varepsilon_0\right] = \mathbb{P}\left[a_1 > b_1 \wedge a_2 > c_1 \mid \varepsilon_a \cap \varepsilon_b\right].$$

Each of those probabilities can be estimated using the following idea: How does the joint distribution of (a_1, a_2) change after conditioning on ε_a?

Let $\tilde{\varphi}_{n-2}(x)$ be the PDF of the distribution of the sum $\sum_{i=3}^{n} a_i/\sqrt{n-2}$. The joint density f_n of (a_1, a_2) conditioned on ε_a must be proportional to $f(a_1) f(a_2)$ multiplied by a "correction factor"

$$\varphi_{n-2}(-a_1 - a_2) := \sqrt{2\pi}\tilde{\varphi}_{n-2}((-a_1 - a_2)/\sqrt{n-2}),$$

which is $\sqrt{2\pi(n-2)}$ times larger than the density of $\sum_{i=1}^{n-2} a_i$ conditioned on ε_a (our normalization is chosen so that $\varphi_{n-2}(x) \approx 1$ for $x \approx 0$):

$$f_n(a_1, a_2) = C_n f(a_1) f(a_2) \varphi_{n-2}(-a_1 - a_2)$$

for some normalization constant $C_n \approx 1$. By the CLT, we should have

$$\varphi_{n-2}(-x) \approx \exp\left(-\frac{x^2}{2(n-2)}\right) \approx 1 - \frac{x^2}{2n}, \tag{2.15}$$

and consequently

$$\mathbb{P}\left[a_1 > b_1 \wedge a_2 > c_1 \mid \varepsilon_a \cap \varepsilon_b\right] \approx C_n C_n' \iint_D f(a_1) f(a_2) f(b_1) f(c_1) \left(1 - \frac{(a_1 + a_2)^2 + b_1^2}{2n}\right) da_1 da_2 db_1 dc_1, \tag{2.16}$$

where $D := \{(a_1, a_2, b_1, c_1) : a_1 > b_1 \wedge a_2 > c_1\}$ and C_n' is another normalization constant corresponding to the one-dimensional "density" $\varphi_{n-1}(-b_1)$. From here, (2.16) can be handled by elementary calculus. The actual computations are more complicated, since we have to carefully track errors, including those introduced by the CLT.

Calculation lemma We will go over the variance computation assuming the following lemma, which will be proved afterwards.

Lemma 2 *Let x be a random variable distributed according to F and let*

$$\begin{aligned} A &:= \mathbb{E}[x \cdot F(x)], \\ B &:= \mathbb{E}[x^2 \cdot F(x)], \\ \alpha_1 &:= \frac{5\gamma_3^2}{24} - \frac{\gamma_4}{8}, \\ \alpha_2 &:= \frac{\gamma_3}{2}, \end{aligned}$$

where γ_j denotes the jth cumulant of x. For $k \in \{a, b, c\}$, denote by ε_k the single-die event $\sum_{i=1}^n k_i = 0$. We have the following expressions:

$$\mathbb{P}[a_1 > b_1 \wedge a_2 > b_2 \mid \varepsilon_0] = \frac{1}{4} - \frac{2A^2}{n} + o(n^{-1}), \tag{2.17}$$

$$\mathbb{P}[a_1 > b_1 \mid \varepsilon_a] = \frac{1}{2} + \frac{1}{4n} + \frac{\alpha_2 A}{n} - \frac{B}{2n} + o(n^{-1}), \tag{2.18}$$

$$\mathbb{P}[a_1 > b_1 \wedge a_2 > b_2 \mid \varepsilon_a] = \frac{1}{4} + \frac{1}{4n} + \frac{\alpha_2 A}{n} - \frac{B}{2n} - \frac{A^2}{n} + o(n^{-1}), \tag{2.19}$$

$$\mathbb{P}[a_1 > b_1 \wedge a_2 > c_1 \mid \varepsilon_a \cap \varepsilon_b] = \frac{1}{4} + \frac{1}{8n} + \frac{\alpha_2 A}{2n} - \frac{B}{4n} - \frac{A^2}{n} + o(n^{-1}). \tag{2.20}$$

Furthermore:

$$\mathbb{P}[a_1 > b_1 \wedge a_1 > b_2 \mid \varepsilon_0] = \frac{1}{3} + o(1), \tag{2.21}$$

$$\mathbb{P}[a_1 > b_1 \wedge a_1 > b_2 \mid \varepsilon_a] = \frac{1}{3} + o(1), \tag{2.22}$$

$$\mathbb{P}[a_1 > b_1 \wedge a_1 > c_1 \mid \varepsilon_a \cap \varepsilon_b] = \frac{1}{3} + o(1). \tag{2.23}$$

Since these expressions might look intimidating, let us point out what we think is one of the most important properties: In contrast to (2.17), it turns out that

$$\mathbb{P}[a_1 > b_1 \wedge a_2 > c_1 \mid \varepsilon_0] = \frac{1}{4} - \frac{A^2}{n} + o(n^{-1}).$$

The fact that the errors of order n^{-1} in those two expressions differ by exactly a factor of two turns out to imply that $W^{(ab)} + W^{(bc)} + W^{(ca)}$ has small variance, which, together with anticoncentration argument for $W^{(ab)}$, implies transitivity similarly as in the proof of Theorem 1. Lemma 2 is more complicated since we are relating random variables $W^{(ab)}$ and $V^{(ab)}$, but the $\frac{A^2}{n}$ terms are still crucial, with other terms canceling out one way or another.

Proof of Proposition 1 assuming Lemma 2 We address each of the three terms in (2.14) in turn. First, using (2.21) and (2.17),

$$\begin{aligned}
&\mathrm{Var}[W \mid \varepsilon_0] = \mathrm{Var}\left[\sum_{i,j=1}^{n} W_{ij} \mid \varepsilon_0\right] \\
&= O(n^2) + 2n^2(n-1)\,\mathrm{Cov}\,[W_{11}, W_{12} \mid \varepsilon_0] + n^2(n-1)^2\,\mathrm{Cov}\,[W_{11}, W_{22} \mid \varepsilon_0] \\
&= O(n^2) + 2n^2(n-1)\left(\mathbb{P}[a_1 > b_1 \wedge a_1 > b_2 \mid \varepsilon_0] - \frac{1}{4}\right) \\
&\quad + n^2(n-1)^2\left(\mathbb{P}[a_1 > b_1 \wedge a_2 > b_2 \mid \varepsilon_0] - \frac{1}{4}\right) \\
&= n^3\left(\frac{1}{6} - 2A^2\right) + o(n^3)\,. \qquad (2.24)
\end{aligned}$$

Second, by (2.22), (2.18) and (2.19),

$$\begin{aligned}
&\mathrm{Var}\left[\sum_{i=1}^{n} F(a_i) - F(b_i) \mid \varepsilon_0\right] = 2\,\mathrm{Var}\left[\sum_{i=1}^{n} F(a_i) \mid \varepsilon_0\right] \\
&= 2n\,\mathrm{Var}[F(a_1) \mid \varepsilon_0] + 2n(n-1)\,\mathrm{Cov}[F(a_1), F(a_2) \mid \varepsilon_0] \\
&= 2n\left(\mathbb{E}\left[F(a_1)^2 \mid \varepsilon_0\right] - \mathbb{E}\,[F(a_1) \mid \varepsilon_0]^2\right) \\
&\quad + 2n(n-1)\left(\mathbb{E}\,[F(a_1)F(a_2) \mid \varepsilon_0] - \mathbb{E}\,[F(a_1) \mid \varepsilon_0]^2\right) \\
&= 2n\left(\mathbb{P}\,[a_1 > b_1 \wedge a_1 > b_2 \mid \varepsilon_a] - \mathbb{P}\,[a_1 > b_1 \mid \varepsilon_a]^2\right) \\
&\quad + 2n(n-1)\left(\mathbb{P}\,[a_1 > b_1 \wedge a_2 > b_2 \mid \varepsilon_a] - \mathbb{P}\,[a_1 > b_1 \mid \varepsilon_a]^2\right) \\
&= n\left(\frac{1}{6} - 2A^2\right) + o(n)\,. \qquad (2.25)
\end{aligned}$$

Finally, recalling F_n is the conditional CDF of a_1 given ε_a, and using (2.23), (2.20) and (2.18) again, we have

$$\begin{aligned}
&\sum_{i,j,k=1}^{n} \mathbb{E}\left[\mathbb{I}(a_i > b_j)\,(F(a_k) - F(b_k)) \mid \varepsilon_0\right] \\
&= \sum_{i,j,k=1}^{n} \mathbb{E}\left[F_n(a_i)F(a_k) - (1 - F_n(b_j))F(b_k) \mid \varepsilon_0\right] \\
&= 2n \sum_{i,j=1}^{n} \mathbb{E}\left[F_n(a_i)F(a_j) \mid \varepsilon_0\right] - n^2 \sum_{i=1}^{n} \mathbb{E}\left[F(a_i) \mid \varepsilon_0\right] \\
&= 2n^2\,\mathbb{E}\left[F_n(a_1)F(a_1) \mid \varepsilon_0\right] + 2n^2(n-1)\,\mathbb{E}\left[F_n(a_1)F(a_2) \mid \varepsilon_0\right] - n^3\,\mathbb{E}\left[F(a_1) \mid \varepsilon_0\right] \\
&= 2n^2\mathbb{P}\left[a_1 > b_1 \wedge a_1 > c_1 \mid \varepsilon_a \cap \varepsilon_b\right] + 2n^2(n-1)\mathbb{P}\left[a_1 > b_1 \wedge a_2 > c_1 \mid \varepsilon_a \cap \varepsilon_b\right] \\
&\qquad - n^3\mathbb{P}\left[a_1 > b_1 \mid \varepsilon_a\right] \\
&= n^2\left(\frac{1}{6} - 2A^2\right) + o(n^2)\,. \qquad (2.26)
\end{aligned}$$

Substituting (2.24), (2.25) and (2.26) into (2.14) gives

$$\mathrm{Var}\left[W - n\sum_{i=1}^{n} F(a_i) - F(b_i)\right] = o(n^3)\,.$$

□

It remains to prove Lemma 2.

Integration lemma The technical part of the proof of Lemma 2 consists of the following lemma that replaces the expressions for φ_{n-2} and φ_{n-1} with an appropriate polynomial approximation. Recall the constants α_1 and α_2 defined in the statement of Lemma 2 and that we defined φ_{n-k} as the PDF of $\sum_{i=1}^{n-k} a_i$ multiplied by $\sqrt{2\pi(n-k)}$.

Lemma 3 *Let D be a measurable set in $\mathbb{R}^4$ and write*

$$f(a,b,c,d) := f(a)f(b)f(c)f(d) \quad \text{and} \quad f(a,b) := f(a)f(b).$$

Setting $a := a_1 + a_2$ and $b := b_1 + b_2$ and denoting Lebesgue integration over $da_1da_2db_1db_2$ by dab, we have

$$\begin{aligned}
&\iint_D f(a_1,a_2,b_1,b_2)\cdot\varphi_{n-2}(-a)\varphi_{n-2}(-b)\,dab \\
&\quad = \iint_D f(a_1,a_2,b_1,b_2)\cdot\left(1 + \frac{2\alpha_1}{n} + \frac{\alpha_2(a+b)}{n} - \frac{a^2+b^2}{2n}\right)dab + o(n^{-1})\,. \qquad (2.27)
\end{aligned}$$

Furthermore, using similar notational conventions, we get, for $a := a_1$ *and* $b := b_1$ *(and* $D \subseteq \mathbb{R}^2$*):*

$$\iint_D f(a,b)\cdot\varphi_{n-1}(-a)\,dab = \iint_D f(a,b)\cdot\left(1+\frac{\alpha_1}{n}+\frac{\alpha_2 a}{n}-\frac{a^2}{2n}\right)dab+o(n^{-1})\,; \tag{2.28}$$

for $a := a_1 + a_2$ *and* $b := b_1 + b_2$*:*

$$\iint_D f(a_1,a_2,b_1,b_2)\cdot\varphi_{n-2}(-a)\,dab = \iint_D f(a_1,a_2,b_1,b_2)\cdot\left(1+\frac{\alpha_1}{n}+\frac{\alpha_2 a}{n}-\frac{a^2}{2n}\right)dab+o(n^{-1})\,; \tag{2.29}$$

and for $a := a_1 + a_2$*,* $b := b_1$ *and* $c := c_1$*:*

$$\iint_D f(a_1,a_2,b,c)\cdot\varphi_{n-2}(-a)\varphi_{n-1}(-b)\,dabc = \iint_D f(a_1,a_2,b,c)\cdot\left(1+\frac{2\alpha_1}{n}+\frac{\alpha_2(a+b)}{n}-\frac{a^2+b^2}{2n}\right)dabc+o(n^{-1})\,. \tag{2.30}$$

We state all formulas that we need explicitly in order to avoid defining and handling new notation, but we point out the pattern in these expressions: the α_1/n factor is multiplied by the number of the densities in the expression, the α_2/n factor is multiplied by the sum of all variables featured in the densities and the quadratic factor is consistent with the approximation (2.15).

Before proving the lemma we point out a corollary that follows by setting D to the full integration space and some simple integration (keeping in mind $\mathbb{E}[a_1] = 0$ and $\mathbb{E}[a_1^2] = 1$). The corollary allows us to estimate the normalization constants C_n and C'_n (see (2.16)).

Corollary 1 *Keeping the notation from Lemma 3, we have*

$$\iint_{\mathbb{R}^4} f(a_1,a_2,b_1,b_2)\cdot\varphi_{n-2}(-a)\varphi_{n-2}(-b)\,dab = 1+\frac{2\alpha_1}{n}-\frac{2}{n}+o(n^{-1}),$$

$$\iint_{\mathbb{R}^2} f(a,b)\cdot\varphi_{n-1}(-a)\,dab = 1+\frac{\alpha_1}{n}-\frac{1}{2n}+o(n^{-1}),$$

$$\iint_{\mathbb{R}^4} f(a_1,a_2,b_1,b_2)\cdot\varphi_{n-2}(-a)\,dab = 1+\frac{\alpha_1}{n}-\frac{1}{n}+o(n^{-1}),$$
$$\iint_{\mathbb{R}^4} f(a_1,a_2,b,c)\cdot\varphi_{n-2}(-a)\varphi_{n-1}(-b)\,dabc = 1+\frac{2\alpha_1}{n}-\frac{3}{2n}+o(n^{-1}).$$

Consequently, letting $D=\{(a_1,a_2,b_1,b_2): a_1>b_1\wedge a_2>b_2\}$, *we have*

$$\begin{aligned}
&\mathbb{P}[a_1>b_1\wedge a_2>b_2\mid\varepsilon_0]\\
&\quad=\frac{\iint_D f(a_1,a_2,b_1,b_2)\varphi_{n-2}(-a)\varphi_{n-2}(-b)\,dab}{\iint_{\mathbb{R}^4} f(a_1,a_2,b_1,b_2)\varphi_{n-2}(-a)\varphi_{n-2}(-b)\,dab}\\
&\quad=\left(1-\frac{2\alpha_1}{n}+\frac{2}{n}\right)\iint_D f(a_1,a_2,b_1,b_2)\Big(1+\frac{2\alpha_1}{n}+\frac{\alpha_2(a+b)}{n}\\
&\qquad-\frac{a^2+b^2}{2n}\Big)\,dab+o(n^{-1})\\
&\quad=\left(1+\frac{2}{n}\right)\iint_D f(a_1,a_2,b_1,b_2)\Big(1+\frac{2\alpha_2(a_1+b_1)}{n}\\
&\qquad\qquad-\frac{a_1^2+b_1^2+a_1a_2+b_1b_2}{n}\Big)\,dab+o(n^{-1}).
\end{aligned}\tag{2.31}$$

Similarly, we have

$$\begin{aligned}
&\mathbb{P}[a_1>b_1\mid\varepsilon_a]\\
&=\left(1-\frac{\alpha_1}{n}+\frac{1}{2n}\right)\iint_D f(a,b)\Big(1+\frac{\alpha_1}{n}+\frac{\alpha_2 a}{n}-\frac{a^2}{2n}\Big)dab+o(n^{-1}),\\
&=\left(1+\frac{1}{2n}\right)\iint_D f(a_1,b_1)\Big(1+\frac{\alpha_2a_1}{n}-\frac{a_1^2}{2n}\Big)dab+o(n^{-1}),
\end{aligned}\tag{2.32}$$

where $D=\{(a_1,b_1): a_1>b_1\}$;

$$\begin{aligned}
&\mathbb{P}[a_1>b_1\wedge a_2>b_2\mid\varepsilon_a]\\
&=\left(1-\frac{\alpha_1}{n}+\frac{1}{n}\right)\iint_D f(a_1,a_2,b_1,b_2)\left(1+\frac{\alpha_1}{n}+\frac{\alpha_2 a}{n}-\frac{a^2}{2n}\right)dab+o(n^{-1})\\
&=\left(1+\frac{1}{n}\right)\iint_D f(a_1,a_2,b_1,b_2)\left(1+\frac{2\alpha_2a_1}{n}-\frac{a_1^2+a_1a_2}{n}\right)dab+o(n^{-1}),
\end{aligned}\tag{2.33}$$

where $D = \{(a_1, a_2, b_1, b_2) : a_1 > b_1 \wedge a_2 > b_2\}$;

$$\begin{aligned}
&\mathbb{P}\left[a_1 > b_1 \wedge a_2 > c_1 \mid \varepsilon_a \cap \varepsilon_b\right] \\
&\quad = \left(1 - \frac{2\alpha_1}{n} + \frac{3}{2n}\right) \iint_D f(a_1, a_2, b, c)\left(1 + \frac{2\alpha_1}{n} + \frac{\alpha_2(a+b)}{n}\right. \\
&\qquad\qquad \left. - \frac{a^2 + b^2}{2n}\right) dabc + o(n^{-1}) \\
&\quad = \left(1 + \frac{3}{2n}\right) \iint_D f(a_1, a_2, b_1, c_1)\left(1 + \frac{\alpha_2(2a_1 + b_1)}{n}\right. \\
&\qquad\qquad \left. - \frac{2a_1^2 + b_1^2 + 2a_1a_2}{2n}\right) dabc + o(n^{-1}), \qquad (2.34)
\end{aligned}$$

where $D = \{(a_1, a_2, b_1, c_1) : a_1 > b_1 \wedge a_2 > c_1\}$.

We point out that an important feature of the expressions (2.31)–(2.34) is that the number of mixed a_1a_2 and b_1b_2 terms depends on the number of φ_{n-2} densities in the expression.

Proof of Lemma 2 assuming Lemma 3 We delay the proof of Lemma 3 and prove Lemma 2 now. For this we need some elementary integral computations. First, in the case with two variables a, b and $D_2 := \{(a, b) : a > b\}$:

$$\begin{aligned}
&\iint_{D_2} f(a,b)\, dab = \frac{1}{2}, \\
&\iint_{D_2} f(a,b) \cdot a\, dab = \int_{-\infty}^{+\infty} af(a) \int_{-\infty}^{a} f(b)\, dbda = \mathbb{E}\left[a \cdot F(a)\right] = A, \\
&\iint_{D_2} f(a,b) \cdot a^2\, dab = \int_{-\infty}^{+\infty} a^2 f(a) \int_{-\infty}^{a} f(b)\, dbda = \mathbb{E}\left[a^2 \cdot F(a)\right] = B.
\end{aligned} \qquad (2.35)$$

In the four-variable case with $D := \{(a_1, a_2, b_1, b_2) : a_1 > b_1 \wedge a_2 > b_2\}$, $f := f(a_1, a_2, b_1, b_2)$ and $dab = da_1da_2db_1db_2$:

$$\begin{aligned}
&\iint_D f\, dab = \frac{1}{4}, \\
&\iint_D f \cdot a_1\, dab = \frac{1}{2} \iint_{D_2} f(a_1, b_1) \cdot a_1\, dab = \frac{A}{2}, \\
&\iint_D f \cdot b_1\, dab = \frac{1}{2} \int_{-\infty}^{+\infty} b_1 f(b_1) \int_{a_1}^{+\infty} f(a_1)\, da_1 db_1 \\
&\qquad = \frac{1}{2} \int_{-\infty}^{+\infty} b_1 f(b_1)(1 - F(b_1))\, db_1 = \frac{\mathbb{E}[b_1] - \mathbb{E}[b_1 \cdot F(b_1)]}{2} = -\frac{A}{2},
\end{aligned}$$

$$\begin{aligned}
\iint_D f \cdot a_1^2 \, dab &= \frac{1}{2} \iint_{D_2} f(a_1, b_1) \cdot a_1^2 \, dab = \frac{B}{2}, \\
\iint_D f \cdot b_1^2 \, dab &= \frac{1}{2} \int_{-\infty}^{+\infty} b_1^2 f(b_1) \int_{b_1}^{+\infty} f(a_1) \, da_1 db_1 = \frac{\mathbb{E}[b_1^2] - \mathbb{E}[b_1^2 \cdot F(b_1)]}{2} \\
&= \frac{1 - B}{2}, \\
\iint_D f \cdot a_1 a_2 \, dab &= \left(\iint_{D_2} f(a_1, b_1) \cdot a_1 \, dab \right)^2 = A^2, \\
\iint_D f \cdot b_1 b_2 \, dab &= \left(\int_{-\infty}^{+\infty} b_1 f(b_1) \int_{b_1}^{+\infty} f(a_1) \, da_1 db_1 \right)^2 = \mathbb{E}[b_1 (1 - F(b_1)]^2 \\
&= A^2. \qquad (2.36)
\end{aligned}$$

Now all that is left is to insert the expressions computed above into Eqs. (2.31)–(2.34) in Corollary 1. For example, in case of (2.34) we get

$$\begin{aligned}
&\mathbb{P}[a_1 > b_1 \wedge a_2 > c_1 \mid \varepsilon_a \cap \varepsilon_b] \\
&\quad = \left(1 + \frac{3}{2n}\right) \iint_D f(a_1, a_2, b_1, c_1) \left(1 + \frac{\alpha_2 (2a_1 + b_1)}{n} \right. \\
&\qquad \left. - \frac{2a_1^2 + b_1^2 + 2a_1 a_2}{2n}\right) dabc + o(n^{-1}) \\
&\quad = \left(1 + \frac{3}{2n}\right) \left(\frac{1}{4} + \frac{\alpha_2 A}{n} - \frac{\alpha_2 A}{2n} - \frac{B}{2n} - \frac{1 - B}{4n} - \frac{A^2}{n}\right) + o(n^{-1}) \\
&\quad = \frac{1}{4} + \frac{1}{8n} + \frac{\alpha_2 A}{2n} - \frac{B}{4n} - \frac{A^2}{n} + o(n^{-1}), \qquad (2.37)
\end{aligned}$$

which is exactly (2.20) that we wanted to prove. Equations (2.17)–(2.19) and (2.21)–(2.23) are handled in analogous ways and we provide the explicit computations only in the "Appendix". □

Proof of Lemma 3 Finally, we turn to Lemma 3. Let $\tilde{\varphi}_j$ denote the density of $j^{-1/2} \sum_{i=1}^{j} a_i$. Since the density of $\sum_{i=1}^{k} a_i$ is bounded for all k (recall that a_i has density continuous on closed support), [30, Theorem 15, pp. 206–207] implies

$$\tilde{\varphi}_j(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2} \left(1 + \frac{\frac{\gamma_3}{3!} H_3(y)}{\sqrt{j}} - \frac{\frac{1}{2}(\frac{\gamma_3}{3!})^2 H_6(y) + \frac{\gamma_4}{4!} H_4(y)}{j}\right) + o(j^{-1}), \qquad (2.38)$$

where γ_j denotes the jth cumulant, the error is uniform in $y \in \mathbb{R}$, and the H_j are Hermite polynomials:

$$\begin{aligned} H_3(y) &= y^3 - 3y, \\ H_4(y) &= y^4 - 6y + 3, \\ H_6(y) &= y^6 - 15y^4 + 45y^2 - 15. \end{aligned} \tag{2.39}$$

Since we defined $\varphi_j(x) = \sqrt{2\pi}\tilde{\varphi}_j(xj^{-1/2})$, (2.38) implies

$$\begin{aligned} \varphi_j(x) &= e^{-x^2/(2j)}\left(1 + \frac{\frac{\gamma_3}{3!}H_3(xj^{-1/2})}{\sqrt{j}} - \frac{\frac{1}{2}(\frac{\gamma_3}{3!})^2 H_6(xj^{-1/2}) + \frac{\gamma_4}{4!}H_4(xj^{-1/2})}{j}\right) \\ &\quad + \mathrm{o}(j^{-1}) \\ &= e^{-x^2/(2j)}\left(1 + \frac{\alpha_1}{j} - \frac{\alpha_2 x}{j}\right) + O\left(\frac{\max(|x|, x^6)}{j^{3/2}}\right) + o(j^{-1}), \end{aligned} \tag{2.40}$$

where in the last line the additional remainder term comes from writing out the Hermite polynomials (2.39) and then noting that what is left out of the main term has smallest order terms $j^{3/2}$ in the denominator, and largest order terms in the numerator x or x^6, depending on $|x| \le 1$ or $|x| > 1$. Substituting this into the left-hand side of (2.27) and using the fact that the sixth moment is finite, we get (letting $f := f(a_1, a_2, b_1, b_2)$)

$$\begin{aligned} &\iint_D f \cdot \varphi_{n-2}(-a)\varphi_{n-2}(-b)\, dab \\ &\quad = \iint_D f \cdot \left[\exp\left(-\frac{a^2}{2(n-2)}\right)\left(1 + \frac{\alpha_1}{n-2} + \frac{\alpha_2 a}{n-2}\right.\right. \\ &\qquad\qquad \left.\left. + O\left(\frac{\max(|a|, a^6)}{n^{3/2}}\right) + o(n^{-1})\right)\right] \\ &\qquad \cdot \left[\exp\left(-\frac{b^2}{2(n-2)}\right)\left(1 + \frac{\alpha_1}{n-2} + \frac{\alpha_2 b}{n-2}\right.\right. \\ &\qquad\qquad \left.\left. + O\left(\frac{\max(|b|, b^6)}{n^{3/2}}\right) + o(n^{-1})\right)\right] dab \\ &\quad = \iint_D f \cdot \exp\left(-\frac{a^2+b^2}{2(n-2)}\right)\left(1 + \frac{2\alpha_1}{n} + \frac{\alpha_2(a+b)}{n}\right) dab + o(n^{-1}) \\ &\quad = \iint_D f \cdot \left(1 - \frac{a^2+b^2}{2(n-2)} + O\left(\min\left(\frac{a^2+b^2}{n}, \frac{(a^2+b^2)^2}{n^2}\right)\right)\right) \\ &\qquad \cdot \left(1 + \frac{2\alpha_1}{n} + \frac{\alpha_2(a+b)}{n}\right) dab + o(n^{-1}) \\ &\quad = \iint_D f \cdot \left[1 + \frac{2\alpha_1}{n} + \frac{\alpha_2(a+b)}{n} - \frac{a^2+b^2}{2n}\right. \\ &\qquad \left. + O\left(\min\left(\frac{a^2+b^2}{n}, \frac{(a^2+b^2)^2}{n^2}\right)\right)\right] dab + o(n^{-1}), \end{aligned} \tag{2.41}$$

where we used the approximation $\exp(-x) = 1 - x + O(\max(x, x^2))$ for $x \ge 0$.

Inspecting (2.41), we see that all that is left to establish (2.27) is to show

$$\iint_{\mathbb{R}^4} f \cdot \left(\min \left(\frac{a^2+b^2}{n}, \frac{(a^2+b^2)^2}{n^2} \right) \right) dab = o(n^{-1}). \tag{2.42}$$

We do that by dividing the integration area into two parts:

$$D_1 := \{(a_1, a_2, b_1, b_2) : a^2 + b^2 < n^{1/3}\} \text{and } D_2 := \mathbb{R}^4 \setminus D_1,$$

and computing

$$\begin{aligned} &\iint_{\mathbb{R}^4} f \cdot \left(\min \left(\frac{a^2+b^2}{n}, \frac{(a^2+b^2)^2}{n^2} \right) \right) dab \\ &\quad \le \iint_{D_1} f \cdot \left(\frac{(a^2+b^2)^2}{n^2} \right) dab + \iint_{D_2} f \cdot \left(\frac{a^2+b^2}{n} \right) dab = O(n^{-4/3}), \end{aligned}$$

where in the inequality we bound the minimum by one of the terms, and then use the fact that small moments of a and b are finite, and that on D_2, we have $1 \le (a^2+b^2)n^{-1/3}$, and therefore $\frac{a^2+b^2}{n} \le \frac{(a^2+b^2)^2}{n^{4/3}}$. Therefore, we have shown (2.27). Similar calculations concerning (2.28)–(2.30) are skipped here and provided in the Appendix. Note that we always need at most sixth finite moment when estimating (2.40). □

3 Stationary Gaussian dice

3.1 Preparation

Before we state and prove our results, let us start with some useful facts about Gaussian Hilbert spaces. It is a well-known fact that the Hermite polynomials $\{H_k, k \ge 0\}$ are *orthogonal* polynomials with respect to the standard Gaussian measure $\gamma(A) = \int_A \varphi(x)\,dx$, for any Borel set $A \subset \mathbb{R}$. Here φ is the standard Gaussian density function and H_k can be defined via *Rodrigues' formula*: $H_k(x) = (-1)^k \varphi(x)^{-1} \frac{d^k}{dx^k}(\varphi(x))$.

For any $f \in L^2(\mathbb{R}, \gamma)$, we have

$$f = \sum_{q \ge 0} \mathrm{coef}(q) H_q \quad \text{with} \mathrm{coef}(q) := \frac{1}{q!} \int_{\mathbb{R}} H_q(x) f(x)\, \gamma(dx)\,,$$

where the above series converges in $L^2(\mathbb{R}, \gamma)$; see [28, Sect. 1.4]. In our work, we only need (1.4) and (1.5). We can find the expansion (1.4), for instance, in [20, p. 7]. Suppose $Z \sim N(0,1)$, noting that $\mathbb{E}\left[(\mathbb{I}[Z > 0] - 2^{-1})^2\right] = 1/4$, we deduce from the orthogonality relation of Hermite polynomials that

$$\frac{1}{4} = \sum_{k \ge 0} d_{2k+1}^2 (2k+1)!\,, \tag{3.1}$$

from which together with the explicit expression of d_{2k+1}'s, we can deduce one of *Srinivasa Ramanujan*'s ingenious identities (in a different form):

$$\pi = \sum_{k\geq 0} \frac{1}{2^{2k-1}(2k+1)} \binom{2k}{k}. \tag{3.2}$$

Ramanujan's identity reads as follows:

$$\frac{\pi}{2} = 1 + \frac{1}{2}\left(\frac{1}{3}\right) + \frac{1\cdot 3}{2\cdot 4}\left(\frac{1}{5}\right) + \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\left(\frac{1}{7}\right) + \cdots;$$

see [33].

To obtain (1.5), note that $\Phi(x) = \mathbb{E}\big(\mathbb{I}[-Z < x]\big)$, then using the expansion (1.4), we get

$$\Phi(x) = \mathbb{E}\big(\mathbb{I}[Z/\sqrt{2} + x/\sqrt{2} > 0]\big) = \frac{1}{2} + \mathbb{E}\left(\sum_{q\geq 0} d_{2q+1} H_{2q+1}\big(Z/\sqrt{2} + x/\sqrt{2}\big)\right),$$
$$= \frac{1}{2} + \mathbb{E}\left(\sum_{q\geq 0} d_{2q+1} \sum_{k=0}^{2q+1} \binom{2q+1}{k} 2^{-q-\frac{1}{2}} H_k(Z) H_{2q+1-k}(x)\right)$$

where we deduce the last equality from the well-known identity: for $a, b \in \mathbb{R}$ satisfying $a^2 + b^2 = 1$, $H_n(ax+by) = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} H_k(x) H_{n-k}(y)$. Note that $\mathbb{E}[H_k(Z)] = 0$ for any $k \geq 1$ and $\mathbb{E}[H_0(Z)] = 1$. Therefore, the expansion (1.5) is established.

Remark 4 Newton's 1676 identity reads as follows: (see [1, p. 228])

$$\frac{\pi}{6} = \arcsin(1/2) = \frac{1}{2} + \frac{1}{2}\cdot\frac{1}{3\cdot 2^3} + \frac{1\cdot 3}{2\cdot 4}\cdot\frac{1}{5\cdot 2^5} + \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\cdot\frac{1}{7\cdot 2^7} + \cdots,$$

which is equivalent to

$$\pi = \sum_{q=0}^{\infty} \frac{3}{(2q+1)2^{4q}} \binom{2q}{q}. \tag{3.3}$$

Using the explicit expression (1.5) for ℓ_{2q+1} and noting that $\Phi(G)$ for standard Gaussian G has distribution that is uniform in $(0, 1)$, we easily check that

$$\frac{1}{6} = \sum_{q=0}^{\infty} (2q+1)!\, 2^{-2q} d_{2q+1}^2, \tag{3.4}$$

from which we have $\alpha = \frac{1}{6} - \frac{1}{2\pi} = \sum_{q=1}^{\infty} (2q+1)!\, 2^{-2q} d_{2q+1}^2$.

Lemma 4 *Suppose X, Y are two centered (jointly) Gaussian random variables with mean zero and variance one such that $\mathbb{E}[XY] = \rho$. Let Φ be the CDF of X, then,*

$$\mathbb{E}\big[\Phi(X)\Phi(Y)\big] = \frac{1}{4} + \sum_{q\geq 0} \ell_{2q+1}^2 (2q+1)!\rho^{2q+1} = \frac{1}{4} + \frac{\rho}{4\pi} + O(\rho^3)$$

where $\ell_{2q+1} = d_{2q+1}2^{-q-\frac{1}{2}} = \dfrac{(-1)^q}{\sqrt{\pi}(2q+1)2^{2q+1}q!}$ *for each integer* $q \geq 0$.

Proof Recall from (1.5) the expansion $\Phi = \frac{1}{2} + \sum_{q\geq 0} \ell_{2q+1} H_{2q+1}$. It is also known (see e.g. Proposition 2.2.1 in [28]) that for $X, Y \sim N(0,1)$ jointly Gaussian and any integers $m, n \geq 0$,

$$\mathbb{E}\big[H_m(X)H_n(Y)\big] = m!\big(\mathbb{E}[XY]\big)^m \delta_{mn}. \tag{3.5}$$

Therefore,

$$\begin{aligned}
&\mathbb{E}\big[\Phi(X)\Phi(Y)\big] \\
&= \frac{1}{4} + \sum_{q\geq 0} \ell_{2q+1}^2 \mathbb{E}\big[H_{2q+1}(X)H_{2q+1}(Y)\big] = \frac{1}{4} + \sum_{q\geq 0} \ell_{2q+1}^2 (2q+1)!\rho^{2q+1} \\
&= \frac{1}{4} + \frac{\rho}{4\pi} + \frac{1}{\pi}\sum_{q\geq 1} \frac{1}{(2q+1)2^{4q+2}} \binom{2q}{q} \rho^{2q+1} \\
&= \frac{1}{4} + \frac{\rho}{4\pi} + O(\rho^3),
\end{aligned}$$

where the last big-O estimate follows from the Newton's identity (3.3). □

3.2 Our results

Now we are in a position to present our results for stationary Gaussian dice. Recall from the introduction that $\{G_i, i \in \mathbb{N}\}$ is a centered stationary Gaussian sequence with the correlation function ρ such that $\rho(0) = 1/2$. Let $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ be i.i.d. copies of $\{G_1, \ldots, G_n\}$, then for $i, j, k, \ell \in [n]$, $(a_i - b_j, a_k - b_\ell)$ is centered bivariate Gaussian with $\mathrm{Var}\left(a_i - b_j\right) = \mathrm{Var}\left(a_k - b_\ell\right) = 1$ and $\mathbb{E}\big[(a_i - b_j)(a_k - b_\ell)\big] = \rho(i-k) + \rho(j-\ell)$. Therefore, we can compute the variance of $W^{(ab)} := \sum_{i,j\in[n]} \mathbb{I}[a_i > b_j]$ using the expansion (1.4) and the relation (3.5):

$$\begin{aligned}
\mathrm{Var}\left(W^{(ab)}\right) &= \sum_{i,j,k,\ell\in[n]} \left\{\mathbb{E}\big(\mathbb{I}[a_i > b_j \wedge a_k > b_\ell]\big) - \frac{1}{4}\right\} \\
&= \sum_{i,j,k,\ell\in[n]} \sum_{q\geq 0} d_{2q+1}^2 (2q+1)!\big(\rho(i-k) + \rho(j-\ell)\big)^{2q+1} \\
&= \sum_{q\geq 0} d_{2q+1}^2 (2q+1)! \sum_{i,j,k,\ell\in[n]} \big(\rho(i-k) + \rho(j-\ell)\big)^{2q+1}.
\end{aligned} \tag{3.6}$$

Let us first look at the almost trivial case where $\rho = s_{1/2}$, that is, when $\rho(i-k) = \frac{1}{2}\delta_{ik}$. In this case, we have by (3.4),

$$\mathrm{Var}\left(W^{(ab)}\right) = \left(\sum_{q\geq 0} d_{2q+1}^2 (2q+1)! \frac{1}{2^{2q}}\right) n^3 + O(n^2) = \frac{1}{6}n^3 + O(n^2). \quad (3.7)$$

Then, by standard computations and the above variance estimate, we have

$$\mathrm{Var}\left(W^{(ab)} - n\sum_{i\in[n]}\left[F(a_i) - F(b_i)\right]\right) = O(n^2),$$

while due to the classical CLT, $n^{-1/2}\sum_{i\in[n]}\left[F(a_i) - F(b_i)\right]$ converges in law to $N(0, 1/6)$. Therefore, we can conclude that the CDF-ordering property (1.3) occurs with high probability in this setting. This relation also implies the following more general result.

Theorem 6 *Let $\boldsymbol{x} = (x_1, \ldots, x_n)$ be a sequence of i.i.d random variables such that x_1 has a density function with a support which is a countable collection of (possibly infinite) intervals. Assume $\boldsymbol{y}$ and $\boldsymbol{z}$ are two i.i.d. copies of $\boldsymbol{x}$, then with high probability,*

$$\boldsymbol{x} \text{ beats } \boldsymbol{y} \text{ if and only if } \sum_{i=1}^n \mathcal{F}(x_i) > \sum_{i=1}^n \mathcal{F}(y_i),$$

where $\mathcal{F}$ is the distribution function (CDF) of x_1. In particular, the probability that $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$ are intransitive tends to zero, as $n \to +\infty$.

Proof Let $\boldsymbol{a}, \boldsymbol{b}$ be given as in the case where $\rho = s_{1/2}$ and F be the distribution function of $a_1 \sim N(0, 1/2)$, then by integral transform, we can assume that

$$\left\{(x_i, y_i) : i \in \mathbb{N}\right\} = \left\{\left(\mathcal{F}^{-1} \circ F(a_i), \mathcal{F}^{-1} \circ F(b_i)\right) : i \in \mathbb{N}\right\},$$

where $\mathcal{F}^{-1}(p) := \inf\{x \in \mathbb{R} : \mathcal{F}(x) \geq p\}$ is the generalized inverse of $\mathcal{F}$. It is clear that $F(a_i) \in (0, 1)$ almost surely and due to our assumption on $\mathcal{F}$, we have $\mathcal{F} \circ \mathcal{F}^{-1}(p) = p$ for any $p \in (0, 1)$. It follows that

$$\begin{aligned}
&\sum_{i=1}^n \mathcal{F}(x_i) > \sum_{i=1}^n \mathcal{F}(y_i) \\
&\overset{\text{with prob. 1}}{\Longleftrightarrow} \quad \sum_{i=1}^n F(a_i) > \sum_{i=1}^n F(b_i) \\
&\overset{\text{with high prob.}}{\Longleftrightarrow} \quad \sum_{i,j=1}^n \mathbb{I}[a_i > b_j] > \frac{n^2}{2} \Leftrightarrow \sum_{i,j=1}^n \mathbb{I}\left[F(a_i) > F(b_j)\right] > \frac{n^2}{2}
\end{aligned}$$

$$\overset{\text{with prob. 1}}{\iff} \quad \sum_{i,j=1}^{n} \mathbb{I}[\mathcal{F}^{-1} \circ F(a_i) > \mathcal{F}^{-1} \circ F(b_j)] > \frac{n^2}{2}$$

$$\iff \quad \sum_{i,j=1}^{n} \mathbb{I}[x_i > y_j] > \frac{n^2}{2}\,.$$

Hence the desired conclusions follow immediately.

□

In the following, we provide the proof of our Theorem 2 as well as some results for the general stationary Gaussian dice. We first state two results of central importance to our approach.

Theorem 7 ([6], Breuer–Major theorem) *Fix an integer $d \geq 1$. Assume $f \in L^2(\mathbb{R}, \gamma)$ admits the following expansion in $L^2(\gamma)$ (Recall $\gamma(dx) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2)dx$):*

$$f = \sum_{q=d}^{\infty} \mathit{coef}(q) H_q \quad \mathit{with} \operatorname{coef}(d) \neq 0; \quad d \text{ is called the Hermite rank of } f\,.$$

Assume also that $(X_k, k \in \mathbb{Z})$ is a centered stationary Gaussian sequence with **unit** *variance*[3] *such that its correlation function $\widetilde{\rho}$ belongs to $\ell^d(\mathbb{Z})$, where $\widetilde{\rho}(i-j) = \mathbb{E}[X_i X_j]$ for any $i, j \in \mathbb{Z}$.*

Then

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} f(X_k) \text{ converges in law to } N(0, \sigma^2) \text{ as } n \to +\infty\,,$$

where $\sigma^2 := \sum_{q=d}^{\infty} q!\, \mathit{coef}(q)^2 \sum_{v \in \mathbb{Z}} \widetilde{\rho}(v)^q \in [0, +\infty)$ is part of the conclusion.

For a modern proof using fourth moment theorems, one can refer to e.g., Theorem 7.2.4 in [28]. In particular, we also need one ingredient from this proof, which we state in the following.

Lemma 5 *Let the assumptions of Theorem 7 be satisfied, that is, $\widetilde{\rho} \in \ell^d(\mathbb{Z})$. For any integer $q \geq d \vee 2$, and any $r \in \{1, \ldots, q-1\}$, we have*

$$n^{-1+\frac{r}{q}} \sum_{|j|<n} |\widetilde{\rho}(j)|^r = o(1) \quad \text{as } n \to +\infty; \quad \text{see equation (7.2.7) in [28].}$$

[3] That is, $\widetilde{\rho}(0) = 1$, which is different from $\rho(0) = 1/2$.

Proof of Theorem 2 Note that we have proved the case where $H = 1/2$. Our proof then consists of only two parts: in the first part, we prove our result for $H \in (1/2, 1)$ and in the second part, we prove a stronger result (Theorem 8) that includes the case $H \in (0, 1/2)$.

We proceed in the same way as in previous subsection: we first estimate the variance of the difference $W^{(ab)} - n\sum_{i=1}^{n}\big[F(a_i) - F(b_i)\big]$, then prove a CLT for $W^{(ab)}$. We begin with the following two lemmas dealing with two variance estimates.

Lemma 6 *Let $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ be i.i.d. copies of the centered stationary Gaussian sequence $\{G_i, i \in \mathbb{N}\}$ with the correlation function ρ such that $\rho(0) = 1/2$. Then*

$$Var\left(W^{(ab)} - n\sum_{i=1}^{n}\big[F(a_i) - F(b_i)\big]\right) = \frac{1}{3}\,Var\left(W^{(ab)} + W^{(bc)} + W^{(ca)}\right) \tag{3.8}$$

$$= \sum_{q\geq 1} d_{2q+1}^2 (2q+1)! \sum_{v=1}^{2q} \binom{2q+1}{v} \left(\sum_{|i|<n} (n-|i|)\rho(i)^v\right) \times \left(\sum_{|j|<n} (n-|j|)\rho(j)^{2q+1-v}\right). \tag{3.9}$$

(1) *If $\rho \in \ell^3(\mathbb{Z})$, then*

$$Var\left(W^{(ab)} - n\sum_{i=1}^{n}\big[F(a_i) - F(b_i)\big]\right) = o(n^3).$$

(2) *Consider $\rho = s_H$, then the case $H \in (0, 5/6)$ is covered by point (1); if $H \in [5/6, 1)$, we have*

$$Var\left(W^{(ab)} - n\sum_{i=1}^{n}\big[F(a_i) - F(b_i)\big]\right) \sim \frac{H^2(2H-1)}{16\pi(4H-3)} n^{6H-2}.$$

The proofs of the above lemma and the following lemma will be postponed to the end of this section.

Lemma 7 *Let $\boldsymbol{a}, \boldsymbol{b}$ and $\{G_i, i \in \mathbb{N}\}$ be given as in Lemma 6. The following statements hold true.*

(1) *If $\rho \in \ell^1(\mathbb{Z})$, then, with $\beta := 2\sum_{q\geq 0} d_{2q+1}^2 (2q+1)! \sum_{i\in\mathbb{Z}} \rho(i)^{2q+1} \in [0, +\infty)$,*

$$Var\left(W^{(ab)}\right) = \beta n^3 + o(n^3).$$

(2) *Consider the case where $\rho = s_H$ is given as in (1.1):*

(i) *for $H \in (0, 1/2]$, $Var\left(W^{(ab)}\right) = \beta n^3 + o(n^3)$ with β defined as in point (1); moreover, $\beta > 0$ in this case.*

(ii) *for* $H \in (1/2, 1)$, $Var\left(W^{(ab)}\right) = \frac{1}{2\pi} n^{2H+2} + o(n^{2H+2})$.

Assuming Lemmas 6 and 7, we prove Theorem 2 in the following. As announced, we split our proof into two cases.
case 1: $H \in (1/2, 1)$. In this case, we deduce from the above two lemmas that

$$\mathrm{Var}\left(W^{(ab)} - n\sum_{i=1}^{n}[F(a_i) - F(b_i)]\right) \Big/ \mathrm{Var}\left(W^{(ab)}\right) = o(1). \tag{3.10}$$

And we have, with $\ell_0 = \frac{1}{2\sqrt{\pi}}$ (see (1.5))

$$\sum_{i\in[n]}\left(F(a_i) - \frac{1}{2}\right) = \sum_{i\in[n]}\left(F(a_i) - \frac{1}{2} - \ell_0\sqrt{2}a_i\right) + \sqrt{2}\ell_0\sum_{i\in[n]} a_i$$

and it is clear that the second part in the above sum is a centered Gaussian with

$$\mathrm{Var}\left(\sqrt{2}\ell_0\sum_{i=1}^{n} a_i\right) = \frac{1}{2\pi}\sum_{i,j=1}^{n} s_H(i-j) \sim \frac{1}{4\pi} n^{2H}, \quad \text{as } n \to +\infty,$$

where the asymptotic behavior is implied by (1.2). We know from (3.10) and point (ii) in Lemma 7 that

$$\mathrm{Var}\left(\sum_{i=1}^{n}\left[F(a_i) - F(b_i)\right]\right) \Big/ \mathrm{Var}\left(\sqrt{2}\ell_0\sum_{i=1}^{n}(a_i - b_i)\right) \xrightarrow{n\to\infty} 1. \tag{3.11}$$

Recall the Slutsky's lemma, which says

$$\text{if } X_n \xrightarrow[n\to\infty]{\text{law}} X \text{ and } Y_n \xrightarrow[n\to\infty]{\text{law}} 0, \text{ then } X_n + Y_n \xrightarrow[n\to\infty]{\text{law}} X.$$

Thus, we deduce from (3.11) and the orthogonality property of Hermite polynomials, $n^{-H}\sum_{i=1}^{n}\left[F(a_i) - F(b_i)\right]$ converges in law to $N\left(0, \frac{1}{2\pi}\right)$, as $n \to +\infty$. Combining (3.10) with Slutsky's lemma again yields

$$\frac{1}{n^{H+1}}\left(W^{(ab)} - \frac{n^2}{2}\right) \xrightarrow[n\to+\infty]{\text{law}} N\left(0, \frac{1}{2\pi}\right).$$

Hence the desired conclusions follow from similar arguments as in the proof of Theorem 1. For the sake of completeness, we sketch it below: first we define $V_n = n\sum_{i=1}^{n}\left(F(a_i) - F(b_i)\right)$, then we have for any $\delta > 0$,

$$\mathbb{P}\left\{\operatorname{sgn}(V_n) \neq \operatorname{sgn}\left(W^{(ab)} - \frac{n^2}{2}\right)\right\}$$
$$\leq \mathbb{P}\left\{\left|\frac{W^{(ab)} - \frac{n^2}{2} - V_n}{n^{H+1}}\right| > \delta\right\} + \mathbb{P}\left\{\left|\frac{W^{(ab)} - \frac{n^2}{2}}{n^{H+1}}\right| \leq \delta\right\},$$

where the lim sup of the RHS, as $n \to +\infty$, is bounded by 2δ. This implies that for $H \in (1/2, 1)$, the relation (1.3) occurs with high probability and thus, the probability of a, b, c being intransitive asymptotically vanishes.
case 2: $H \in (0, 1/2)$. In this case, the correlation function $s_H \in \ell^1(\mathbb{Z})$ and by Lemma 7, $\beta = 2\sum_{q\geq 0} d_{2q+1}^2 (2q+1)! \sum_{i\in\mathbb{Z}} s_H(i)^{2q+1} \in (0, +\infty)$. Then, *case 2* is an immediate consequence of the following theorem.

Theorem 8 *Let* $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ *be i.i.d. copies of* $\{G_1, \ldots, G_n\}$ *with correlation function* $\rho \in \ell^1(\mathbb{Z})$ *such that the constant* β *defined in Lemma 7 is* **strictly** *positive. Then, with high probability,*

$$\sum_{i,j=1}^{n} \mathbb{I}[a_i > b_j] > \frac{n^2}{2} \quad \textit{if and only if} \quad \sum_{i=1}^{n} F(a_i) > \sum_{i=1}^{n} F(b_i)\,, \tag{3.12}$$

where $F(x) = \Phi(\sqrt{2}x)$ *is the distributional function of* $G_1 \sim N(0, 1/2)$. *As a consequence, the probability of three dice* $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ *being intransitive tends to zero, as* $n \to +\infty$.

Proof (*Proof of Theorem 8*) Let us first summarize what we have so far, concerning this proof:

- $\operatorname{Var}\left(W^{(ab)}\right) = \beta n^3 + o(n^3)$, with $\beta \in (0, +\infty)$; see Lemma 7.
- $\operatorname{Var}\left(W^{(ab)} - n\sum_{i=1}^{n}\left[F(a_i) - F(b_i)\right]\right) = o(n^3)$; see Lemma 6.

Putting $X_i = \sqrt{2}a_i$ for each $i \in \mathbb{N}$ and $\widetilde{\rho} = 2\rho$, we apply Theorem 7 for $d = 1$, $f = \Phi - 1/2 = \sum_{q\geq 0} \ell_{2q+1} H_{2q+1}$ and we obtain the following CLT:

$$\frac{1}{\sqrt{n}}\sum_{k=1}^{n}\left(F(a_k) - \frac{1}{2}\right) = \frac{1}{\sqrt{n}}\sum_{k=1}^{n} f(X_k) \xrightarrow[n\to+\infty]{\text{law}} N(0, \beta/2),$$

where the limiting variance, due to Breuer–Major's theorem, should be

$$\sum_{q=0}^{\infty}(2q+1)!\ell_{2q+1}^2 \sum_{v\in\mathbb{Z}}(2\rho(v))^{2q+1}\,,$$

which is indeed equal to $\beta/2$ because of $d_{2q+1}^2 = \ell_{2q+1}^2 2^{2q+1}$ for each integer $q \geq 0$.

Thus, we deduce from the above CLT and Slutsky's lemma that

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big[F(a_i)-F(b_i)\big]\xrightarrow[n\to+\infty]{\text{law}}N(0,\beta)\quad\text{and}\quad\frac{W^{(ab)}-\frac{n^2}{2}}{n^{3/2}}\xrightarrow[n\to+\infty]{\text{law}}N(0,\beta).$$

Hence the desired conclusions follow from the same arguments as in the ending paragraph of *case 1*.

□

To conclude this section, it remains to prove Lemmas 6 and 7. One may have noticed that we haven't used the relation (3.8) in the above proofs. In fact, the relation (3.8) and the following Lemma 8 together imply the point (1) in Lemma 6, and besides the independent interest of such a relation, its proof contains some ingredients for our proof of Lemma 6.

Lemma 8 *Let $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ be i.i.d. copies of $\{G_i, i\in\mathbb{N}\}$. Assume that $\rho\in\ell^3(\mathbb{Z})$, then*

$$Var\Big(W^{(ab)}+W^{(bc)}+W^{(ca)}\Big)=o(n^3). \tag{3.13}$$

Proof Using Hermite expansion of $x\in\mathbb{R}\longmapsto\mathbb{I}[x>0]$, we have

$$\begin{aligned}W^{(ab)}+W^{(bc)}+W^{(ca)}&=\sum_{i,j=1}^{n}\big(\mathbb{I}[a_i>b_j]+\mathbb{I}[b_i>c_j]+\mathbb{I}[c_i>a_j]\big)\\&=\frac{3n^2}{2}+\sum_{q\geq0}d_{2q+1}\sum_{i,j=1}^{n}\Big[H_{2q+1}(a_i-b_j)+H_{2q+1}(b_i-c_j)+H_{2q+1}(c_i-a_j)\Big]\end{aligned}$$

so that

$$\begin{aligned}\mathrm{Var}\Big(W^{(ab)}+W^{(bc)}+W^{(ca)}\Big)&=\sum_{q\geq0}d_{2q+1}^2(2q+1)!\\&\times\sum_{i,j,k,\ell=1}^{n}\Big(\mathbb{E}[(a_i-b_j)(a_k-b_\ell)]^{2q+1}+\mathbb{E}[(a_i-b_j)(b_k-c_\ell)]^{2q+1}\\&+\mathbb{E}[(a_i-b_j)(c_k-a_\ell)]^{2q+1}+\mathbb{E}[(b_i-c_j)(a_k-b_\ell)]^{2q+1}+\mathbb{E}[(b_i-c_j)(b_k-c_\ell)]^{2q+1}\\&+\mathbb{E}[(b_i-c_j)(c_k-a_\ell)]^{2q+1}+\mathbb{E}[(c_i-a_j)(a_k-b_\ell)]^{2q+1}\\&+\mathbb{E}[(c_i-a_j)(b_k-c_\ell)]^{2q+1}+\mathbb{E}[(c_i-a_j)(c_k-a_\ell)]^{2q+1}\Big).\end{aligned}$$

Then, using the specific correlation structure of a,b,c as well as their independence, we get

$$\frac{1}{3}\operatorname{Var}\left(W^{(ab)}+W^{(bc)}+W^{(ca)}\right)$$
$$=\sum_{q\geq 1} d_{2q+1}^2(2q+1)!\sum_{i,j,k,\ell=1}^{n}\Big[\big(\rho(i-k)+\rho(j-\ell)\big)^{2q+1}-\rho(i-k)^{2q+1}$$
$$-\rho(j-\ell)^{2q+1}\Big]. \tag{3.14}$$

Let us now look at the second sum in (3.14), which can be rewritten using the binomial formula, as follows:

$$\sum_{i,j,k,\ell=1}^{n}\sum_{v=1}^{2q}\binom{2q+1}{v}\rho(i-k)^v\rho(j-\ell)^{2q+1-v}$$
$$=\sum_{v=1}^{2q}\binom{2q+1}{v}\left(\sum_{i,k=1}^{n}\rho(i-k)^v\right)\left(\sum_{j,\ell=1}^{n}\rho(j-\ell)^{2q+1-v}\right)$$
$$=\sum_{v=1}^{2q}\binom{2q+1}{v}2^{-1-2q}\left(\sum_{|i|<n}(n-|i|)\widetilde{\rho}(i)^v\right)\left(\sum_{|j|<n}(n-|j|)\widetilde{\rho}(j)^{2q+1-v}\right) \tag{3.15}$$

by putting $\widetilde{\rho}=2\rho$. It is clear that the term 2^{-1-2q} will compensate the term $\sum_{v=1}^{2q}\binom{2q+1}{v}$ above. Therefore, we only need the following rough estimate: for $q\geq 1$

$$\sum_{i,j,k,\ell=1}^{n}\Big[\big(\rho(i-k)+\rho(j-\ell)\big)^{2q+1}-\rho(i-k)^{2q+1}-\rho(j-\ell)^{2q+1}\Big]$$
$$=O\left\{n^2\left(\sum_{|i|<n}|\widetilde{\rho}(i)|\right)\left(\sum_{|i|<n}|\widetilde{\rho}(i)|^2\right)\right\},$$

implying

$$\operatorname{Var}\left(W^{(ab)}+W^{(bc)}+W^{(ca)}\right)=O\left\{n^2\left(\sum_{|i|<n}|\widetilde{\rho}(i)|\right)\left(\sum_{|i|<n}|\widetilde{\rho}(i)|^2\right)\right\}.$$

The desired estimate (3.13) follows from Lemma 5 and the assumption $\widetilde{\rho}\in\ell^3(\mathbb{Z})$. □

Proof (*Proof of Lemma 6*)
As in previous variance calculations,

we have

$$\begin{aligned}
&\operatorname{Var}\left(W^{(ab)} - n\sum_{i=1}^{n}\big[F(a_i) - F(b_i)\big]\right)\\
&= \operatorname{Var}\big(W^{(ab)}\big) + \frac{1}{2}n^4 - 2n^2\sum_{i,j=1}^{n}\mathbb{E}\big[F(a_i)F(a_j)\big]\\
&= \operatorname{Var}\big(W^{(ab)}\big) + \frac{1}{2}n^4 - 2n^2\left(\frac{1}{3}n + 2\sum_{1\le i<j\le n}\mathbb{E}\big[F(a_i)F(a_j)\big]\right). \qquad (3.16)
\end{aligned}$$

It follows from Lemma 4 that for $i \neq j$, (also due to $d_{2q+1}^2 = \ell_{2q+1}^2 2^{2q+1}$)

$$\begin{aligned}
\mathbb{E}\big[F(a_i)F(a_j)\big] &= \mathbb{E}\big[\Phi(\sqrt{2}a_i)\Phi(\sqrt{2}a_j)\big]\\
&= \frac{1}{4} + \sum_{q\ge 0} d_{2q+1}^2(2q+1)!\rho(i-j)^{2q+1}. \qquad (3.17)
\end{aligned}$$

Therefore, it is routine to verify using (3.6), (3.16), (3.17), (3.15) and (3.14) that

$$\begin{aligned}
&\operatorname{Var}\left(W^{(ab)} - n\sum_{i=1}^{n}\big[F(a_i) - F(b_i)\big]\right)\\
&= \sum_{q\ge 1} d_{2q+1}^2(2q+1)!\sum_{v=1}^{2q}\binom{2q+1}{v}\left(\sum_{|i|<n}(n-|i|)\rho(i)^v\right)\\
&\quad\times\left(\sum_{|j|<n}(n-|j|)\rho(j)^{2q+1-v}\right) = \frac{1}{3}\operatorname{Var}\big(W^{(ab)} + W^{(bc)} + W^{(ca)}\big). \qquad (3.18)
\end{aligned}$$

Therefore, the relations (3.8) and (3.9) are established. If $\rho \in \ell^3(\mathbb{Z})$, Lemma 8 implies that the variance in (3.18) is $o(n^3)$.

To prove point (2), we consider the particular case where $\rho = s_H$. One can easily verify using the asymptotic relation (1.2) that $s_H \in \ell^3(\mathbb{Z})$ if and only if $H \in (0, 5/6)$. Now suppose that $H \in [5/6, 1)$, the relation (3.9) still holds true, that is, we have

$$\begin{aligned}
&\operatorname{Var}\left(W^{(ab)} - n\sum_{i=1}^{n}\big[F(a_i) - F(b_i)\big]\right)\\
&= \frac{1}{2\pi}\left(\sum_{|i|<n}(n-|i|)s_H(i)\right)\left(\sum_{|j|<n}(n-|j|)s_H(j)^2\right)
\end{aligned}$$

pagebreak

$$+\sum_{q\geq 2} d_{2q+1}^2 (2q+1)! \sum_{v=1}^{2q} \binom{2q+1}{v} \left(\sum_{|i|<n} (n-|i|) s_H(i)^v\right) \times \left(\sum_{|j|<n} (n-|j|) s_H(j)^{2q+1-v}\right).$$

One can readily check using (1.2) that for $H \in [5/6, 1)$,

$$\sum_{|i|<n} \big(n-|i|\big) s_H(i) \sim \frac{1}{2} n^{2H} \quad \text{and} \quad \sum_{|i|<n} \big(n-|i|\big) s_H(i)^2 \sim \frac{H^2(2H-1)}{4(4H-3)} n^{4H-2},$$

and

$$\sum_{|i|<n} \big(n-|i|\big) s_H(i)^3 \sim \begin{cases} \dfrac{H^3(2H-1)^3}{8(6H-5)(3H-2)} n^{6H-4} & \text{if } H \in (5/6, 1) \\ 2(5/18)^3 n \log n & \text{if } H = 5/6. \end{cases}$$

All these estimates imply, whenever $H \in [5/6, 1)$,

$$\operatorname{Var}\left(W^{(ab)} - n\sum_{i=1}^{n} [F(a_i) - F(b_i)]\right) = \frac{1}{3} \operatorname{Var}\left(W^{(ab)} + W^{(bc)} + W^{(ca)}\right) \sim \frac{H^2(2H-1)}{16\pi(4H-3)} n^{6H-2}.$$

Hence the proof of Lemma 6 is completed.

□

Proof (*Proof of Lemma 7*) Assume first that $\rho \in \ell^1(\mathbb{Z})$ and recall from (3.6) that

$$\operatorname{Var}\big(W^{(ab)}\big) = \sum_{q\geq 0} d_{2q+1}^2 (2q+1)! \sum_{i,j,k,\ell=1}^{n} \big(\rho(i-k) + \rho(j-\ell)\big)^{2q+1}$$

and in view of (3.14), we have

$$\operatorname{Var}\big(W^{(ab)}\big) = 2\sum_{q\geq 0} d_{2q+1}^2 (2q+1)! \sum_{i,j,k,\ell=1}^{n} \rho(i-k)^{2q+1} + o(n^3). \tag{3.19}$$

The second sum in (3.19) is equal to $n^2 \sum_{|i|<n} (n-|i|)\rho(i)^{2q+1}$. Since $\rho \in \ell^1(\mathbb{Z})$ and for $q \geq 0$,

$$\lim_{n\to+\infty} \sum_{|i|<n} \frac{n-|i|}{n} \rho(i)^{2q+1} = \sum_{i\in\mathbb{Z}} \rho(i)^{2q+1} \quad \text{by } \textit{dominatedconvergence}.$$

Therefore, as $n \to +\infty$,

$$n^{-3} \sum_{q\geq 0} d_{2q+1}^2 (2q+1)! \sum_{i,j,k,\ell=1}^{n} \rho(i-k)^{2q+1}$$
$$= \sum_{q\geq 0} d_{2q+1}^2 (2q+1)! \sum_{|i|<n} \frac{n-|i|}{n} \rho(i)^{2q+1} \to \sum_{q\geq 0} d_{2q+1}^2 (2q+1)! \sum_{i\in\mathbb{Z}} \rho(i)^{2q+1},$$

so that $\mathrm{Var}\left(W^{(ab)}\right) = \beta n^3 + o(n^3)$. Note that $\beta \in [0, +\infty)$ under the assumption $\rho \in \ell^1(\mathbb{Z})$ is an easy consequence of Theorem 7. It is clear that $\widetilde{\rho} = 2\rho$ satisfies the assumption of Theorem 7, then using $d_{2q+1}^2 = \ell_{2q+1}^2 2^{2q+1}$, we get

$$\frac{1}{2}\beta = \sum_{q\geq 0} d_{2q+1}^2 (2q+1)! 2^{-1-2q} \sum_{i\in\mathbb{Z}} \widetilde{\rho}(i)^{2q+1} = \sum_{q\geq 0} \ell_{2q+1}^2 (2q+1)! \sum_{i\in\mathbb{Z}} \widetilde{\rho}(i)^{2q+1}.$$

So, with $f(x) = \Phi(x) - \frac{1}{2}$ and $d = 1$, one can see that $\beta \in [0, +\infty)$.

Now let us look at the fractional case, and note that the case $H = 1/2$ was stated in (3.7).

If $H < 1/2$, then s_H is summable so that $\sum_{i\in\mathbb{Z}} s_H(i)$ is finite, which is the limit of

$$\sum_{|k|\leq n} s_H(k) = \frac{1}{4} \sum_{|k|\leq n} \left(|k+1|^{2H} + |k-1|^{2H} - 2|k|^{2H}\right) = \frac{1}{2}\left(|n+1|^{2H} - |n|^{2H}\right)$$

as $n \to +\infty$. This limit is zero. For later reference, we summarize some basic properties of s_H for $H \in (0, 1/2)$:

$$s_H(0) = \frac{1}{2}, -\frac{1}{2} < s_H(v) < 0 \text{ for } v \neq 0; \text{ and } \sum_{v\in\mathbb{Z}} s_H(v) = 0. \tag{3.20}$$

It follows that for $q \geq 1$, from

$$1 = 2s_H(0) = \sum_{v\neq 0} \left[-2s_H(v)\right] > \sum_{v\neq 0} \left[-2s_H(v)\right]^{2q+1}$$

we obtain $\sum_{i\in\mathbb{Z}} s_H(i)^{2q+1} \in (0, +\infty)$. Thus, point (2)-(i) is proved.

If $H \in (1/2, 1)$, then $s_H(v) > 0$. One can verify by using (3.16), (3.17) and the fact $1/6 = \sum_{q=0}^{\infty} d_{2q+1}^2 (2q+1)! 2^{-2q}$ from Remark 4 that

$$\mathrm{Var}\left(W^{(ab)}\right) = \mathrm{Var}\left(W^{(ab)} - n\sum_{i=1}^{n}\left[F(a_i) - F(b_i)\right]\right)$$
$$+ 2n^2 \sum_{q\geq 0} d_{2q+1}^2 (2q+1)! \sum_{i,j\in[n]} s_H(i-j)^{2q+1}.$$

The first term in the above sum is of order $o(n^{2H+2})$, by Lemma 6. It remains to use (1.2) to estimate the second term in the above sum:

$$2n^2 d_0^2 \sum_{i,j\in[n]} s_H(i-j) \sim \frac{1}{2\pi} n^{2H+2}$$

gives the dominant contribution. Hence our proof of Lemma 7 is now completed. □

4 Condorcet paradox for close elections: majority

This section contains the proof of Theorem 3.

4.1 Notation

We start with recalling and extending the model and notation. There are n voters (where n is odd) and each of them independently chooses one of $k!$ rankings of the alternatives uniformly at random. For voter i, such a random ranking gives rise to a random tuple $x_i = (x_i^{(1)}, \ldots, x_i^{(K)})$ in $\{-1, 1\}^K$ representing $K := \binom{k}{2}$ pairwise choices (according to some fixed ordering of pairs). We call each of $k!$ tuples in the support of x_i *transitive*. Any other tuple is *intransitive*. We say that a tuple has a *Condorcet winner* if it has an alternative that beats everyone else.

We denote aggregation over voters by boldface. Therefore, we write $\mathbf{x} = (x_1, \ldots x_n)$ for the random vector of voter preferences (where each element is itself a random tuple of length K).

For $j = 1, \ldots, K$, let $S_i^{(j)} := \sum_{i'=1}^{i} x_{i'}^{(j)}$ and $S^{(j)} := S_n^{(j)}$, and write

$$Y^{(j)} = \mathrm{Maj}_n(\mathbf{x}^{(j)}) = \mathrm{sgn}(S^{(j)}) \,.$$

Furthermore, we write $Y = \left(Y^{(1)}, \ldots, Y^{(K)}\right)$ and $S = \left(S^{(1)}, \ldots, S^{(K)}\right)$ for the aggregated tuples.

Given voter preferences, we say that the voting outcome is intransitive if the aggregated tuple Y is intransitive. Similarly, we say that there is a Condorcet winner if tuple Y has a Condorcet winner.

We are interested in situations where elections are "almost tied" or, more precisely, "d-close" for $d \geq 1$. Specifically, we define ε_d to be the event where $\|S\|_\infty \leq d$, i.e., $|S^{(j)}|$ is at most d for every $j \in [K]$.

4.2 Local CLT

We use a theorem and some definitions from the textbook on random walks by Spitzer [34]. In accordance with the book, we make

Definition 1 A k-dimensional random walk $(X_i)_{i\in\mathbb{N}}$ is a Markov chain over $\mathbb{Z}^k$ with $X_0 = 0^k$ and a distribution of one step $Z_{i+1} := X_{i+1} - X_i$ that does not depend on i.

Defining $S_i := (S_i^{(1)}, \ldots, S_i^{(K)})$, note that $(S_i)_{i\in\{0,\ldots,n\}}$ is a K-dimensional random walk and that we want to calculate $\mathbb{P}(\mathrm{sgn}(S_n) = y|\varepsilon_d)$, for $y \in \{-1, 1\}^K$. There is one technicality we need to address to apply a local CLT: since the steps of our random walk are in $\{-1, 1\}^K$, the values of (S_i) lie on a proper sublattice of $\mathbb{Z}^K$, namely, $S_i^{(j)}$ always has the same parity as i. To deal with this, we define $T_i^{(j)} := (S_{2i+1}^{(j)} - 1)/2$. Note that (T_i) is still a K-dimensional random walk, with one catch: the starting point T_0 is not necessarily the origin, but rather one of $k!$ points in $\{-1, 0\}^K$ corresponding to the transitive tuple picked by the first voter.

Before we state the local CLT, we need another definition:

Definition 2 [34, D1 in Sect. 5] A K-dimensional random walk is *strongly aperiodic* if for every $t \in \mathbb{Z}^K$, the subgroup of $\mathbb{Z}^K$ generated by the points that can be reached from t in one step is equal to $\mathbb{Z}^K$.

Now we are ready to state the theorem:

Theorem 9 (Local CLT, Remark after P9 in Sect. 7 of [34]) *Let $(T_i)_{i\in\mathbb{N}}$ be a strongly aperiodic K-dimensional random walk, starting at origin and with a single step Z, i.e., $T_{i+1} - T_i$ distributed according to Z.*

If $\mathbb{E}[Z] = 0^K$ and Q is the $K \times K$ (finite) covariance matrix of Z, then matrix Q is invertible and for every $t \in \mathbb{Z}^K$,

$$\left|(2\pi n)^{K/2}\,\mathbb{P}\left[T_n = t\right] - |Q|^{-1/2}\exp\left(\frac{-t^T Q^{-1} t}{2n}\right)\right| = o(1)\ ,$$

where the $o(1)$ function depends on n, but not on t.

Our main lemma states that the distribution of T_n conditioned on $\|T_n\|_\infty$ being small is roughly uniform.

Lemma 9 *For the random walk (T_i) defined above and $t \in \mathbb{Z}^K, d \geq 1$ such that $\|t\|_\infty \leq d$, there are some $\alpha_k, \beta_k > 0$ such that*

$$\left|\alpha_k n^{K/2}\mathbb{P}\left[T_n = t\right] - 1\right| \leq \beta_k \frac{d^2}{n} + o_k(1)\ . \tag{4.1}$$

Proof We first deal with the technicality that we mentioned before: the starting point T_0 of the random walk is itself a random variable. In the proof below we proceed by conditioning on $T_0 = 0^K$. After reading the proof it should be clear how to modify it for other starting points in $\{-1, 0\}^K$. Equation (4.1) is obtained from those conditional results by triangle inequality.

We need to check that the random walk (T_i) satisfies hypothesis of Theorem 9. First, note that the "step" random variable Z for (T_i) has the same distribution as $(X_1 + X_2)/2$, i.e., two steps of our original random process.

Clearly, $\mathbb{E}[Z] = (\mathbb{E}[X_1] + \mathbb{E}[X_2])/2 = 0^K$. Equally clearly, all covariances in the matrix Q are finite.

To show that (T_i) is strongly aperiodic, let $(e^{(1)}, \ldots, e^{(K)})$ be the standard basis of $\mathbb{Z}^K$. Note that it is enough to show that for each $z \in \mathbb{Z}^K$, all of $z, z+e^{(1)}, \ldots, z+e^{(K)}$ are reachable from z in one step. But this is so:

- It is possible to stay at z by choosing a permutation (ranking) τ for X_1 and then its reverse τ^R for X_2.
- We explain how one can move from z to $z + e^{(j)}$ on an example and hope it is clear how to generalize it. For $k = 5$ and $e^{(j)}$ corresponding to the b versus d comparison, one can choose a ranking $b > d > a > c > e$ for X_1 followed by $e > c > a > b > d$ for X_2.

Since Theorem 9 applies, we have

$$\left|(2\pi n)^{K/2}\,\mathbb{P}\,[T_n = t] - |Q|^{-1/2}\exp\left(-t^T Q^{-1} t/2n\right)\right| = o_k(1),$$

which can be rewritten as

$$\left|\alpha_k n^{K/2}\mathbb{P}\,[T_n = t] - \exp\left(-t^T Q^{-1} t/2n\right)\right| = o_k(1)\ .$$

Since $1 - x \le \exp(-x) \le 1$ for $x \ge 0$, it follows that

$$\left|\alpha_k n^{K/2}\mathbb{P}\,[T_n = t] - 1\right| \le \frac{t^T Q^{-1} t}{2n} + o_k(1)\ .$$

Finally we observe that $t = dt'$ for some t' with $\|t'\|_\infty \le 1$, so we have

$$\frac{t^T Q^{-1} t}{2n} \le \beta_k \frac{d^2}{n}\ ,$$

as we needed. □

Lemma 9 implies:

Corollary 2 *Let n be odd, $d \ge 1$ and $s \in (2\mathbb{Z}+1)^K$ be a tuple such that $\|s\|_\infty \le d$. Then for some $\alpha_k, \beta_k > 0$,*

$$\left|\alpha_k\,(n-1)^{K/2}\,\mathbb{P}\,[S = s] - 1\right| \le \beta_k \frac{d^2}{n} + o_k(1)\ .$$

Proof Letting $t := (s - 1^K)/2$, note that $\mathbb{P}[S_n = s] = \mathbb{P}\left[T_{(n-1)/2} = t\right]$ and that $\|t\|_\infty \le d$. We get the result by applying Lemma 9. □

4.3 Proof of Theorem 3

Recall that we want to prove (1.8), that is

$$\left|\mathbb{P}\left[Y = y \mid \varepsilon_d\right] - \frac{1}{2^K}\right| \leq \alpha_k \frac{d^2}{n} + o(1) \,.$$

After we have (1.8), the bounds (1.9) and (1.10) easily follow by triangle inequality.

For $y \in \{-1,1\}^K$, let $\mathcal{S}_y := \left\{s \in (2\mathbb{Z}+1)^K : \bigwedge_{j\in[K]} \operatorname{sgn}\left(s^{(j)}\right) = y^{(j)} \wedge \|s\|_\infty \leq d\right\}$. Observe that $\mathbb{P}[Y = y \wedge \varepsilon_d] = \sum_{s\in\mathcal{S}_y} \mathbb{P}[S = s]$. Furthermore, note that $|\mathcal{S}_y| = |\mathcal{S}_{y'}|$ for every y, y'. Set $M := |\mathcal{S}_y|$ as the common cardinality of the $\mathcal{S}_y$ sets.

First, we use Corollary 2 to show that the probability $\mathbb{P}[Y = y \mid \varepsilon_d]$ must be close to $q := \frac{1}{\alpha_k (n-1)^{K/2}} \cdot \frac{M}{\mathbb{P}[\varepsilon_d]}$, where α_k is the constant from Corollary 2:

$$\begin{aligned}
\left|\frac{\mathbb{P}[Y = y \mid \varepsilon_d]}{q} - 1\right| &= \left|\frac{\alpha_k (n-1)^{K/2}\mathbb{P}[\varepsilon_d]}{M} \cdot \mathbb{P}[Y = y \mid \varepsilon_d] - 1\right| \\
&= \left|\frac{\alpha_k (n-1)^{K/2}}{M} \cdot \sum_{s\in\mathcal{S}_y} \mathbb{P}[S = s] - 1\right| \\
&\leq \frac{1}{M} \sum_{s\in\mathcal{S}_y} \left|\alpha_k (n-1)^{K/2}\mathbb{P}[S = s] - 1\right| \leq \beta_k \frac{d^2}{n} + o(1) \,.
\end{aligned}$$

The value of q depends on k, n and d, but not on y. The implication is that the conditional probabilities must be almost equal for every pair y, y':

$$\begin{aligned}
\left|\mathbb{P}[Y = y \mid \varepsilon_d] - \mathbb{P}[Y = y' \mid \varepsilon_d]\right| &\leq \left|\mathbb{P}[Y = y \mid \varepsilon_d] - q\right| + \left|q - \mathbb{P}[Y = y' \mid \varepsilon_d]\right| \\
&\leq 2q\left(\beta_k \frac{d^2}{n} + o(1)\right) \leq \beta'_k \frac{d^2}{n} + o(1) \,.
\end{aligned}$$

But this is all we need, since

$$\begin{aligned}
\left|\mathbb{P}[Y = y \mid \varepsilon_d] - \frac{1}{2^K}\right| &\leq \frac{1}{2^K} \sum_{y'\in\{-1,1\}^K} \left|\mathbb{P}[Y = y \mid \varepsilon_d] - \mathbb{P}[Y = y' \mid \varepsilon_d]\right| \\
&\leq \beta_k \frac{d^2}{n} + o(1) \,.
\end{aligned}$$

□

Remark 5 A similar bound with an explicit $o(1)$ term of the order $O_k\left(\frac{d}{\sqrt{n}}\right) + O_k\left(\frac{n^{K/2-1}}{d^K}\right)$ (implying chaotic behavior for $n^{1/2-1/K} \ll d \ll n^{1/2}$) can be achieved using the multidimensional Berry–Esseen theorem instead of the local CLT.

Remark 6 As we mentioned in Sect. 1.3, the proof of Theorem 3 can be modified to give a similar bound

$$\mathbb{P}\left[Y = y \mid \varepsilon_d^{(a_0b_0)}\right] = \frac{1}{2^K} + o(1)$$

for $d = o(\sqrt{n})$ also in case the event $\varepsilon_d^{(a_0b_0)}$ is defined as $\left|S^{(ab)}\right| \le d$ for all pairwise comparisons (ab) different from (a_0b_0).

The reason for this is that if we remove conditioning from just one $S^{(a_0b_0)}$, there are still no covariance factors in the CLT computation that would steer the distribution of Y away from uniform.

5 Condorcet paradox for close elections: majority of triplets

Recall that we are considering odd $n = 3m$ voters, alternatives a, b, c and random variables $x_1^{(kk')}, \ldots, x_n^{(kk')}$ and that the pairwise comparison is done according to $f\colon \{-1,1\}^n \to \{-1,1\}$:

$$f(x_1, \ldots, x_n) = \operatorname{sgn}\left(\sum_{i=1}^{m} \operatorname{sgn}(w_i)\right), \quad \text{where } w_i = x_{3i-2} + x_{3i-1} + x_{3i}.$$

This section contains proofs of non-chaotic behavior of f under certain conditionings. Section 5.1 contains the proof of Theorem 4, dealing with conditioning on small $\left|\sum_{i=1}^n x_i^{(kk')}\right|$. In Sect. 5.2 we prove Theorem 5, which considers conditioning on small $\left|T_\rho f(x^{(kk')})\right|$.

5.1 Proof of Theorem 4

For $i \in [m]$, we take random tuple $Z_i := \left(A_i^{(kk')}, B_i^{(kk')}\right)_{(kk')}$ for $kk' \in \{ab, bc, ca\}$, where $A_i^{(kk')} := w_i^{(kk')}/\sqrt{3}$ and $B_i^{(kk')} := \operatorname{sgn}\left(w_i^{(kk')}\right)$. Note that $Z_1, \ldots, Z_m$ are i.i.d. Let us compute the first two moments of the single-voter distribution $Z = (A^{(ab)}, A^{(bc)}, A^{(ca)}, B^{(ab)}, B^{(bc)}, B^{(ca)})$. For this keep in mind that $\operatorname{Cov}\left[x_i^{(kk')}, x_i^{(k'k'')}\right] = -1/3$ and refer to Table 1 for the joint distribution of $w^{(kk')}$ and $w^{(k'k'')}$:

$$\begin{aligned}
\mathbb{E}\left[A^{(kk')}\right] &= \mathbb{E}\left[B^{(kk')}\right] = 0\\
\operatorname{Var}\left[A^{(kk')}\right] &= \operatorname{Var}\left[B^{(kk')}\right] = 1\\
\operatorname{Cov}\left[A^{(kk')}, A^{(k'k'')}\right] &= -\frac{1}{3}\\
\operatorname{Cov}\left[B^{(kk')}, B^{(k'k'')}\right] &= \frac{80-136}{8\cdot 27} = -\frac{7}{27}
\end{aligned}$$

Table 1 Probabilities of values for $w^{(kk')}, w^{(k'k'')}$ pairs multiplied by common denominator $8 \cdot 27$

$w^{(kk')}$ versus $w^{(k'k'')}$	-3	-1	1	3
-3	1	6	12	8
-1	6	27	36	12
1	12	36	27	6
3	8	12	6	1

Keep in mind that $x_i^{(kk')}$ and $x_i^{(k'k'')} \in \{-1, 1\}$ are equal with probability 1/3

$$\begin{aligned}
\mathrm{Cov}\left[A^{(kk')}, B^{(kk')}\right] &= \frac{1}{\sqrt{3}} \cdot \frac{3}{2} = \frac{\sqrt{3}}{2} \\
\mathrm{Cov}\left[A^{(kk')}, B^{(k'k'')}\right] &= \frac{1}{\sqrt{3}} \cdot \frac{3 \cdot 14 + 66 - 96 - 3 \cdot 40}{8 \cdot 27} = -\frac{1}{2\sqrt{3}}\,. \qquad (5.1)
\end{aligned}$$

Let $\tilde{A}^{(kk')} := \sum_{i=1}^m A_i^{(kk')}/\sqrt{m}$ and $\tilde{B}^{(kk')} := \sum_{i=1}^m B_i^{(kk')}/\sqrt{m}$ and let $\tilde{M}^{(kk')}$ and $\tilde{N}^{(kk')}$ be joint standard Gaussians with the same covariance structure as $\tilde{A}^{(kk')}$ and $\tilde{B}^{(kk')}$ respectively. After checking that our six by six covariance matrix is not singular, by the multi-dimensional Berry–Esseen theorem (see the statement e.g., in [4]), we can move to the Gaussian space:

$$\begin{aligned}
&\mathbb{P}\left[f(\mathbf{x}^{(ab)}) = f(\mathbf{x}^{(bc)}) = f(\mathbf{x}^{(ca)}) \wedge \varepsilon_d\right] \\
&\quad = 2\mathbb{P}\left[f(\mathbf{x}^{(ab)}) = f(\mathbf{x}^{(bc)}) = f(\mathbf{x}^{(ca)}) = 1 \wedge \varepsilon_d\right] \\
&\quad = 2\mathbb{P}\left[\|\tilde{A}\|_\infty \le \frac{d}{\sqrt{3m}} \wedge \tilde{B} \ge 0\right] \\
&\quad = 2\mathbb{P}\left[\|\tilde{M}\|_\infty \le \frac{1}{\log n} \wedge \tilde{N} \ge 0\right] + O\left(\frac{1}{\sqrt{n}}\right), \qquad (5.2)
\end{aligned}$$

where we write $\tilde{B} \ge 0$ to indicate $\tilde{B}^{(kk')} \ge 0$ for every component of $\tilde{B}$. Similarly,

$$\mathbb{P}[\varepsilon_d] = \mathbb{P}\left[\|\tilde{M}\|_\infty \le \frac{1}{\log n}\right] + O\left(\frac{1}{\sqrt{n}}\right).$$

Let us define three more centered Gaussians $\tilde{R}^{(kk')}$ according to the formula

$$\tilde{N}^{(kk')} = \frac{\sqrt{3}}{2}\tilde{M}^{(kk')} + \frac{1}{2}\tilde{R}^{(kk')}. \qquad (5.3)$$

Since $\mathrm{Cov}[\tilde{M}^{(kk')}, \tilde{N}^{(kk')}] = \mathrm{Cov}[A^{(kk')}, B^{(kk')}] = \sqrt{3}/2$, we immediately see that $\mathrm{Var}[\tilde{R}^{(kk')}] = 1$ and $\mathrm{Cov}[\tilde{M}^{(kk')}, \tilde{R}^{(kk')}] = 0$. Furthermore, we calculate

$$\mathrm{Cov}[\tilde{M}^{(kk')}, \tilde{R}^{(k'k'')}] = 2\,\mathrm{Cov}[\tilde{M}^{(kk')}, \tilde{N}^{(k'k'')}] - \sqrt{3}\,\mathrm{Cov}[\tilde{M}^{(kk')}, \tilde{M}^{(k'k'')}]$$

$$= 2\,\mathrm{Cov}[A^{(kk')}, B^{(k'k'')}] - \sqrt{3}\,\mathrm{Cov}[A^{(kk')}, A^{(k'k'')}] = 0,$$

$$\mathrm{Cov}[\tilde{R}^{(kk')}, \tilde{R}^{(k'k'')}] = 4\,\mathrm{Cov}[\tilde{N}^{(kk')}, \tilde{N}^{(k'k'')}] - 4\sqrt{3}\,\mathrm{Cov}[\tilde{M}^{(kk')}, \tilde{N}^{(k'k'')}]$$
$$+ 3\,\mathrm{Cov}[\tilde{M}^{(kk')}, \tilde{M}^{(k'k'')}] = -\frac{1}{27}. \tag{5.4}$$

Recall the joint density function for centered Gaussians: in k dimensions, for the distribution with covariance matrix Σ and $\mathbf{x} = (x_1, \ldots, x_k)$ we have

$$f_\Sigma(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^k |\Sigma|}} \exp\left(-\mathbf{x}^T \Sigma^{-1} \mathbf{x}\right).$$

In particular, letting $c_\Sigma := f_\Sigma(0)$, we have basic approximation

$$f_\Sigma(\mathbf{x}) = c_\Sigma + O(\|\mathbf{x}\|^2). \tag{5.5}$$

Letting $D := \{\boldsymbol{m} \in \mathbb{R}^3 : \|\boldsymbol{m}\|_\infty \le 1/\log n\}$ and using this approximation, we have

$$\mathbb{P}\left[\|\tilde{M}\|_\infty \le \frac{1}{\log n}\right] = \int_D f_M(\boldsymbol{m})\,d\boldsymbol{m} = \frac{8c_M}{\log^3 n} + O\left(\frac{1}{\log^5 n}\right).$$

As for calculating (5.2), given $\boldsymbol{m} \in D$, let

$$D_{\boldsymbol{m}} := \left\{\boldsymbol{r} \in \mathbb{R}^3 : \frac{\sqrt{3}}{2}\boldsymbol{m} + \frac{1}{2}\boldsymbol{r} \ge 0\right\}.$$

In particular, we have $D_0 = \{\boldsymbol{r} : \boldsymbol{r} \ge 0\}$. Let f_R be the density function of the Gaussian triple $\tilde{R}$ and let

$$\alpha^* := 2\mathbb{P}[\tilde{R} \ge 0] = 2\int_{D_0} f_R(\boldsymbol{r})\,d\boldsymbol{r}.$$

Note that if $\|\boldsymbol{m}\|_\infty \le 1/\log n$ and $\boldsymbol{r} \in D_0 \Delta D_{\boldsymbol{m}}$, then there exists at least one coordinate on which $|r_i| = O(1/\log n)$. Therefore, we obtain

$$\left|\int_{D_{\boldsymbol{m}}} f_R(\boldsymbol{r})\,d\boldsymbol{r} - \frac{\alpha^*}{2}\right| \le \int_{D_0 \Delta D_{\boldsymbol{m}}} f_R(\boldsymbol{r})\,d\boldsymbol{r}$$
$$\le 3\mathbb{P}\left[|\tilde{R}^{(kk')}| \le O\left(\frac{1}{\log n}\right)\right] = O\left(\frac{1}{\log n}\right),$$

where the error term is uniform in $\boldsymbol{m}$.

Finally, we recall (5.4) to observe that Gaussian triples $\tilde{M}$ and $\tilde{R}$ are independent and therefore their joint density decomposes $f_{M,R}(\boldsymbol{m}, \boldsymbol{r}) = f_M(\boldsymbol{m}) f_R(\boldsymbol{r})$. That allows us to calculate, using (5.3),

$$\mathbb{P}\left[\|\tilde{M}\|_\infty \le \frac{1}{\log n} \wedge \tilde{N} \ge 0\right] = \int_D f_M(\boldsymbol{m}) \int_{D_m} f_R(\boldsymbol{r})\,d\boldsymbol{r}d\boldsymbol{m}$$

$$= \int_D f_M(\boldsymbol{m}) \left(\frac{\alpha^*}{2} + O\left(\frac{1}{\log n}\right)\right) d\boldsymbol{m} = \frac{8c_M}{\log^3 n} \cdot \frac{\alpha^*}{2} + O\left(\frac{1}{\log^4 n}\right).$$

In conclusion, we get

$$\begin{aligned}
&\mathbb{P}\left[f(\mathbf{x}^{(ab)}) = f(\mathbf{x}^{(bc)}) = f(\mathbf{x}^{(ca)}) \mid \varepsilon_d \right] \\
&\qquad = \frac{2\mathbb{P}\left[\|\tilde{M}\|_\infty \le 1/\log n \wedge \tilde{N} \ge 0 \right] + O(1/\sqrt{n})}{\mathbb{P}\left[\|\tilde{M}\|_\infty \le 1/\log n \right] + O(1/\sqrt{n})} \\
&\qquad = \frac{\frac{8c_M}{\log^3 n}\alpha^* + O(1/\log^4 n)}{\frac{8c_M}{\log^3 n} + O(1/\log^5 n)} = \alpha^* + O\left(\frac{1}{\log n}\right) \\
&\qquad\qquad \xrightarrow{n\to\infty} \alpha^* \approx 23.2\%,
\end{aligned}$$

where in the very last step we employed a computer algebra system to compute the approximate value of α^*.

5.2 Proof of Theorem 5

The proof of Theorem 5 is a refinement of the proof of Theorem 4, which is a recommended preliminary reading. In particular, we will use the notation that was developed there. From now on the constants in the $O(\cdot)$ notation are allowed to depend on ρ. Recall that for $\mathbf{x} \in \{-1, 1\}^n$ and $\mathbf{w} \in \{\pm 3, \pm 1\}^m$ we have defined

$$\begin{aligned}
W_b(\mathbf{x}) = W_b(\mathbf{w}) &= |\{i \in [m] : w_i = b\}| \ , \\
V_b(\mathbf{x}) = V_b(\mathbf{w}) &= W_b(\mathbf{w}) - \mathbb{E}_{\mathbf{w'}}[W_b(\mathbf{w'})] = W_b(\mathbf{w}) - \begin{cases} n/8 & \text{if } b = \pm 3 \ , \\ 3n/8 & \text{if } b = \pm 1 \ . \end{cases}
\end{aligned}$$

We can write $W_b(\mathbf{w}) = \sum_{i=1}^m W_b(w_i)$ and $V_b(\mathbf{w}) = \sum_{i=1}^m V_b(w_i)$ in an obvious way, with $W_b(w_i) \in \{0, 1\}$, $V_{\pm 3}(w_i) \in \{-1/8, 7/8\}$ and $V_{\pm 1}(w_i) \in \{-3/8, 5/8\}$. Note that $W_3(w_i) + W_1(w_i) + W_{-1}(w_i) + W_{-3}(w_i) = 1$ and $V_3(w_i) + V_1(w_i) + V_{-1}(w_i) + V_{-3}(w_i) = 0$.

Taking $w_i = x_{3i-2} + x_{3i-1} + x_{3i}$, $w'_i = x'_{3i-2} + x'_{3i-1} + x'_{3i}$, $s_i = \operatorname{sgn}(w_i)$ and $s'_i = \operatorname{sgn}(w'_i)$, where (x_i, x'_i) are ρ-correlated, we also define

$$\varepsilon := \mathbb{P}\left[x_j \ne x'_j \right] = (1 - \rho)/2 \ , \tag{5.6}$$

$$p_3 := \mathbb{P}\left[s_i = s'_i \mid w_i = 3 \right] = (1 - \varepsilon)^3 + 3\varepsilon(1 - \varepsilon)^2 \ , \tag{5.7}$$

$$p_1 := \mathbb{P}\left[s_i = s'_i \mid w_i = 1 \right] = (1 - \varepsilon)^3 + \varepsilon(1 - \varepsilon)^2 + 2\varepsilon^2(1 - \varepsilon) \ . \tag{5.8}$$

Recall that

$$T_\rho f(\mathbf{x}) = \mathbb{E}_{\mathbf{x'}\sim N_\rho(\mathbf{x})}[f(\mathbf{x'})]$$

and observe that for our particular function f the value of $T_\rho f$ depends only on $\mathbf{w}$ and equals

$$T_\rho f(\mathbf{w}) = \mathbb{E}_{\mathbf{s'}\sim N_\rho(\mathbf{w})}\left[\operatorname{sgn}\left(\sum_{i=1}^m s_i'\right)\right] = 2\mathbb{P}\left[\sum_{i=1}^m s_i' > 0\right] - 1,$$

where random variables $s_i' \in \{-1, 1\}$ are independent and $\mathbb{P}[s_i = s_i'] = p_b$ if $|w_i| = b$ for $b = 1, 3$. In particular, we can also write $T_\rho f(\mathbf{w})$ as a sum of four independent binomial random variables

$$\begin{aligned} T_\rho f(\mathbf{w}) = 2\mathbb{P}\Big[&\operatorname{Bin}(W_3(\mathbf{w}), p_3) + \operatorname{Bin}(W_1(\mathbf{w}), p_1) \\ &+ \operatorname{Bin}(W_{-1}(\mathbf{w}), 1-p_1) + \operatorname{Bin}(W_{-3}(\mathbf{w}), 1-p_3) > \frac{m}{2}\Big] - 1\,. \end{aligned} \tag{5.9}$$

Our plan is to use a CLT argument to conclude that, for most values of $\mathbf{w}$ under event $\mathcal{F}_{\rho,d}$, the value of $T_\rho f(\mathbf{w})$ is proportional to

$$\begin{aligned} T_\rho f(\mathbf{w}) &\asymp \frac{p_3 W_3(\mathbf{w}) + p_1 W_1(\mathbf{w}) + (1-p_1) W_{-1}(\mathbf{w}) + (1-p_3) W_{-3}(\mathbf{w}) - m/2}{\sqrt{m}} \\ &= \frac{p_3 V_3(\mathbf{w}) + p_1 V_1(\mathbf{w}) + (1-p_1) V_{-1}(\mathbf{w}) + (1-p_3) V_{-3}(\mathbf{w})}{\sqrt{m}} \\ &= \frac{q_3 V_3(\mathbf{w}) + q_1 V_1(\mathbf{w}) - q_1 V_{-1}(\mathbf{w}) - q_3 V_{-3}(\mathbf{w})}{\sqrt{m}}\,, \end{aligned}$$

where $q_3 := p_3 - 1/2$ and $q_1 := p_1 - 1/2$. We now state this more precisely as a lemma, the proof of which we defer until later:

Lemma 10 *Let* $\sigma_3^2 := p_3(1-p_3)$, $\sigma_1^2 := p_1(1-p_1)$ *and* $\sigma^2 := \frac{\sigma_3^2 + 3\sigma_1^2}{4}$. *Let*

$$\begin{aligned} A_i^{(kk')} &:= q_3 V_3\big(w_i^{(kk')}\big) + q_1 V_1\big(w_i^{(kk')}\big) - q_1 V_{-1}\big(w_i^{(kk')}\big) - q_3 V_{-3}\big(w_i^{(kk')}\big)\,, \\ \tilde{A}^{(kk')} &:= \frac{1}{\sqrt{m}} \sum_{i=1}^m A_i^{(kk')}\,. \end{aligned}$$

Take $C := \sqrt{\frac{\pi}{2}}\sigma$ *and define events*

$$\begin{aligned} \mathcal{G}_1 :\equiv \mathcal{F}_{\rho,d} \equiv &\quad \max\Big(\big|T_\rho f(\boldsymbol{x}^{(ab)})\big|, \big|T_\rho f(\boldsymbol{x}^{(bc)})\big|, \big|T_\rho f(\boldsymbol{x}^{(ca)})\big|\Big) \le \frac{1}{\log m}\,, \\ \mathcal{G}_2 :\equiv &\quad \|\tilde{A}\|_\infty = \max\Big(\big|\tilde{A}^{(ab)}\big|, \big|\tilde{A}^{(bc)}\big|, \big|\tilde{A}^{(ca)}\big|\Big) \le \frac{C}{\log m}\,. \end{aligned}$$

Let Δ stand for the symmetric difference of events. Then,

$$\mathbb{P}\left[\mathcal{G}_1 \Delta \mathcal{G}_2\right] \leq O\left(\frac{1}{\log^5 m}\right) .$$

Assuming Lemma 10 we continue along the lines of the proof of Theorem 4, letting $B_i^{(kk')} := \operatorname{sgn}(w_i^{(kk')})$ and $Z_i := \left(A_i^{(kk')}, B_i^{(kk')}\right)_{(kk')}$. The random variables $Z_1, \ldots, Z_m$ are i.i.d. and for CLT purposes we can compute (again Table 1 is helpful) the six by six covariance matrix of the distribution of $Z := Z_1$:

$$\begin{aligned}
\mathbb{E}\left[A^{(kk')}\right] &= \mathbb{E}\left[B^{(kk')}\right] = 0 \\
\operatorname{Var}\left[A^{(kk')}\right] &= \frac{q_3^2 + 3q_1^2}{4} \\
\operatorname{Var}\left[B^{(kk')}\right] &= 1 \\
\operatorname{Cov}\left[A^{(kk')}, A^{(k'k'')}\right] &= \frac{-14q_3^2 - 24q_1q_3 - 18q_1^2}{216} \qquad (5.10) \\
\operatorname{Cov}\left[B^{(kk')}, B^{(k'k'')}\right] &= \frac{80 - 136}{8 \cdot 27} = -\frac{7}{27} \\
\operatorname{Cov}\left[A^{(kk')}, B^{(kk')}\right] &= \frac{q_3 + 3q_1}{4} \\
\operatorname{Cov}\left[A^{(kk')}, B^{(k'k'')}\right] &= \frac{-26q_3 - 30q_1}{216} \qquad (5.11)
\end{aligned}$$

Let $\left(\tilde{M}^{(kk')}, \tilde{N}^{(kk')}\right)_{(kk')}$ be joint Gaussians with the same covariance structure as $\left(\tilde{A}^{(kk')}, \tilde{B}^{(kk')}\right)_{(kk')}$. Further symbolic computations in a computer algebra system lead to expressing $\tilde{N}^{(kk')}$ as a linear combination

$$\tilde{N}^{(kk')} = \beta \tilde{M}^{(kk')} + \beta' \left(\tilde{M}^{(k'k'')} + \tilde{M}^{(k''k)}\right) + \gamma \tilde{R}^{(kk')} , \qquad (5.12)$$

where $\gamma > 0$, random tuples $\left(\tilde{M}^{(kk')}\right)_{(kk')}$ and $\left(\tilde{R}^{(kk')}\right)_{(kk')}$ are independent of each other and each $\tilde{R}^{(kk')}$ is a standard Gaussian. Furthermore, we obtain

$$\operatorname{Cov}\left[\tilde{R}^{(kk')}, \tilde{R}^{(k'k'')}\right] = \operatorname{Cov}(\rho) \qquad (5.13)$$

with $\operatorname{Cov}(\rho)$ a decreasing function of $\rho \in (0, 1)$ and

$$\operatorname{Cov}(\rho) \leq -\frac{1}{27} = \lim_{\rho \to 0^+} \operatorname{Cov}(\rho).$$

Since the mutual covariance $\mathrm{Cov}(\rho)$ is decreasing, the expression

$$\alpha(\rho) := 2\mathbb{P}[\tilde{R} \geq 0]$$

is also decreasing in ρ, with $\lim_{\rho\to 0^+} \alpha(\rho) = \alpha^*$ and $\alpha(\rho) \geq \lim_{\rho\to 1^-} \alpha(\rho) \geq 0.17$.

Let $\delta := C/\log m$ and recall Lemma 10. We apply this lemma and similar arguments as in the proof of Theorem 4 and calculate

$$\begin{aligned}
&\mathbb{P}\left[f\left(\mathbf{x}^{(ab)}\right) = f\left(\mathbf{x}^{(bc)}\right) = f\left(\mathbf{x}^{(ca)}\right) \mid \mathcal{F}_{\rho,d}\right] \\
&\quad = \frac{2\mathbb{P}\left[f\left(\mathbf{x}^{(ab)}\right) = f\left(\mathbf{x}^{(bc)}\right) = f\left(\mathbf{x}^{(ca)}\right) = 1 \wedge \mathcal{G}_1\right]}{\mathbb{P}[\mathcal{G}_1]} \\
&\quad = \frac{2\mathbb{P}\left[\tilde{B} \geq 0 \wedge \mathcal{G}_2\right] + O(1/\log^5 m)}{\mathbb{P}[\mathcal{G}_2] + O(1/\log^5 m)} \\
&\quad = \frac{2\mathbb{P}\left[\tilde{B} \geq 0 \wedge \|\tilde{A}\|_\infty \leq \delta\right] + O(1/\log^5 m)}{\mathbb{P}\left[\|\tilde{A}\|_\infty \leq \delta\right] + O(1/\log^5 m)} \\
&\quad = \frac{2\mathbb{P}\left[\tilde{N} \geq 0 \wedge \|\tilde{M}\|_\infty \leq \delta\right] + O(1/\log^5 m)}{\mathbb{P}\left[\|\tilde{M}\|_\infty \leq \delta\right] + O(1/\log^5 m)} \\
&\quad = \frac{8c_M\delta^3 \cdot \alpha(\rho) + O(1/\log^4 m)}{8c_M\delta^3 + O(1/\log^5 m)} = \alpha(\rho) + O\left(\frac{1}{\log m}\right).
\end{aligned} \tag{5.14}$$

It remains to prove Lemma 10.

Proof (*Proof of Lemma 10*) Recall the definitions of $W_b(\mathbf{w})$ and $V_b(\mathbf{w})$. We begin with estimating $T_\rho f(\mathbf{w})$ for a fixed $\mathbf{w}$. In the following we will sometimes drop dependence on $\mathbf{w}$ (writing, e.g., W_b, V_b, $\tilde{A}$ instead of $W_b(\mathbf{w})$, $V_b(\mathbf{w})$, $\tilde{A}(\mathbf{w})$) in the interest of clarity. Recall equation (5.9) and let $Z := \sum_{i=1}^m Z_i$ be the sum of m independent random variables arising out of the four binomial distributions featured there. We have:

$$\begin{aligned}
T_\rho f(\mathbf{w}) &= 2\mathbb{P}\left[Z > \frac{m}{2}\right] - 1\,, \\
\mathbb{E}\left[Z - \frac{m}{2}\right] &= p_3W_3 + p_1W_1 + (1-p_1)W_{-1} + (1-p_3)W_{-3} - \frac{m}{2} \\
&= p_3V_3 + p_1V_1 + (1-p_1)V_{-1} + (1-p_3)V_{-3} \\
&= q_3V_3 + q_1V_1 - q_1V_{-1} - q_3V_{-3} = \sqrt{m}\tilde{A}\,, \\
\mathrm{Var}[Z] &= \sigma_3^2(W_3 + W_{-3}) + \sigma_1^2(W_1 + W_{-1}) \\
&= m\sigma^2 + \sigma_3^2(V_3 + V_{-3}) + \sigma_1^2(V_1 + V_{-1}) = m\sigma^2(1+t)\,,
\end{aligned}$$

for $t := t(\mathbf{w}) := \frac{\sigma_3^2(V_3+V_{-3})+\sigma_1^2(V_1+V_{-1})}{\sigma^2 m}$. Since random variables Z_i are bounded, we can apply the Berry–Esseen theorem and, using $\mathrm{erf}(x/\sqrt{2}) = 2\Phi(x) - 1$ where

$\operatorname{erf}(y) := \frac{2}{\sqrt{\pi}} \int_0^y e^{-s^2} ds$, find

$$\begin{aligned}
\mathbb{P}\left[Z - \frac{m}{2} > 0\right] &= \mathbb{P}\left[\frac{Z - m/2 - \sqrt{m}\tilde{A}}{\sqrt{m}\sigma\sqrt{1+t}} > \frac{-\tilde{A}}{\sigma\sqrt{1+t}}\right] \\
&= \Phi\left(\frac{\tilde{A}}{\sigma\sqrt{1+t}}\right) + O\left(\frac{1}{\sqrt{m(1+t)^3}}\right), \\
T_\rho f(\mathbf{w}) &= \operatorname{erf}\left(\frac{\tilde{A}}{\sqrt{2}\sigma\sqrt{1+t}}\right) + O\left(\frac{1}{\sqrt{m(1+t)^3}}\right). \qquad (5.15)
\end{aligned}$$

From now on we consider a random election with vote vectors $\mathbf{x}^{(ab)}$, $\mathbf{x}^{(bc)}$, $\mathbf{x}^{(ca)}$ that induce $\mathbf{w}^{(ab)}$, $\mathbf{w}^{(bc)}$, $\mathbf{w}^{(ca)}$. First, consider the marginal distribution of $\mathbf{w}$. Since $t(\mathbf{w})$ can be written as a sum of m i.i.d. random variables $\sigma^2 m t(\mathbf{w}) = \sum_{i=1}^m t_i(w_i)$ with $\mathbb{E}[t_i] = 0$ and $|t_i| \le 1$, a standard concentration bound gives

$$\mathbb{P}\left[|t(\mathbf{w})| > \frac{1}{m^{1/4}}\right] \le 2\exp\left(-\frac{\sqrt{m}\sigma^4}{2}\right) \le O\left(\frac{1}{\sqrt{m}}\right). \qquad (5.16)$$

As a consequence of (5.15) and (5.16) and the Taylor expansion $\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}}x + O(x^3)$, whenever $|t| \le m^{-1/4}$ holds, we have

$$T_\rho f(\mathbf{w}) = \frac{\tilde{A}}{C} + O(\tilde{A}^3) + O\left(\frac{1}{m^{1/4}}\right) \qquad (5.17)$$

and, furthermore,

$$|T_\rho f(\mathbf{w})| \le \frac{1}{\log m} \implies |\tilde{A}| \le \frac{C}{\log m} + O\left(\frac{1}{\log^3 m}\right), \qquad (5.18)$$

$$\pm T_\rho f(\mathbf{w}) > \frac{1}{\log m} \implies \pm\tilde{A} \ge \frac{C}{\log m} - O\left(\frac{1}{\log^3 m}\right). \qquad (5.19)$$

We are now ready to bound the measure of the symmetric difference

$$\mathbb{P}[\mathcal{G}_1 \Delta \mathcal{G}_2] = \mathbb{P}[\mathcal{G}_1 \wedge \neg\mathcal{G}_2] + \mathbb{P}[\neg\mathcal{G}_1 \wedge \mathcal{G}_2].$$

We will use the union bound over a small number of cases and show that each of them has probability $O(\log^{-5} m)$.

First, if $\mathcal{G}_1$ holds, but $\mathcal{G}_2$ does not, then $|\tilde{A}^{(kk')}| > C/\log m$ for some comparison (kk'). Let us assume that $\tilde{A}^{(ab)} > C/\log m$, other five cases being symmetrical. We now apply (5.16), (5.18) and multivariate Berry–Esseen and get

$$
\begin{aligned}
&\mathbb{P}\left[\mathcal{G}_1 \wedge \tilde{A}^{(ab)} > \frac{C}{\log m}\right] \\
&\quad \le \mathbb{P}\left[\|\tilde{A}\|_\infty \le \frac{C}{\log m} + O\left(\frac{1}{\log^3 m}\right) \wedge \tilde{A}^{(ab)} > \frac{C}{\log m}\right] \\
&\qquad + \mathbb{P}\left[\|t\|_\infty > \frac{1}{m^{1/4}}\right] \\
&\quad = \mathbb{P}\left[\frac{C}{\log m} \le \tilde{A}^{(ab)} \le \frac{C}{\log m} + O\left(\frac{1}{\log^3 m}\right) \wedge |\tilde{A}^{(bc)}|, |\tilde{A}^{(ca)}| \le \frac{C}{\log m}\right] \\
&\qquad + O\left(\frac{1}{\sqrt{m}}\right) \\
&\quad = \mathbb{P}\left[\frac{C}{\log m} \le \tilde{M}^{(ab)} \le \frac{C}{\log m} + O\left(\frac{1}{\log^3 m}\right) \wedge |\tilde{M}^{(bc)}|, |\tilde{M}^{(ca)}| \le \frac{C}{\log m}\right] \\
&\qquad + O\left(\frac{1}{\sqrt{m}}\right) = O\left(\frac{1}{\log^5 m}\right).
\end{aligned}
$$

Applying union bound over remaining, symmetric cases, we obtain

$$
\mathbb{P}[\mathcal{G}_1 \wedge \neg\mathcal{G}_2] \le O\left(\frac{1}{\log^5 m}\right).
$$

On the other hand, if $\mathcal{G}_2$ holds, but $\mathcal{G}_1$ does not, then we have $|T_\rho f(\mathbf{x}^{(kk')})| > 1/\log m$ for some (kk'), for example, $T_\rho f(\mathbf{x}^{(ab)}) > 1/\log m$. A similar calculation using (5.19) gives

$$
\begin{aligned}
&\mathbb{P}\left[T_\rho f(\mathbf{x}^{(ab)}) > \frac{1}{\log m} \wedge \mathcal{G}_2\right] \\
&\quad \le \mathbb{P}\left[\tilde{A}^{(ab)} \ge \frac{C}{\log m} - O\left(\frac{1}{\log^3 m}\right) \wedge \|\tilde{A}\|_\infty \le \frac{C}{\log m}\right] \\
&\qquad + \mathbb{P}\left[\|t\|_\infty > \frac{1}{m^{1/4}}\right] \\
&\quad = \mathbb{P}\left[\frac{C}{\log m} - O\left(\frac{1}{\log^3 m}\right) \le \tilde{A}^{(ab)} \le \frac{C}{\log m} \wedge |\tilde{A}^{(bc)}|, |\tilde{A}^{(ca)}| \le \frac{C}{\log m}\right] \\
&\qquad + O\left(\frac{1}{\sqrt{m}}\right) \\
&\quad = \mathbb{P}\left[\frac{C}{\log m} - O\left(\frac{1}{\log^3 m}\right) \le \tilde{M}^{(ab)} \le \frac{C}{\log m} \wedge |\tilde{M}^{(bc)}|, |\tilde{M}^{(ca)}| \le \frac{C}{\log m}\right] \\
&\qquad + O\left(\frac{1}{\sqrt{m}}\right) = O\left(\frac{1}{\log^5 m}\right)
\end{aligned}
$$

and

$$\mathbb{P}[\neg\mathcal{G}_1 \wedge \mathcal{G}_2] = O\left(\frac{1}{\log^5 m}\right).$$

□

6 Arrow's theorem for dice

Arguably the most famous result in social choice theory is Arrow's impossibility theorem [2,3]. Intuitively, it states that the only reasonable voting systems based on pairwise comparisons that never produce a Condorcet paradox are "dictators", i.e., functions whose value depend only on a single voter.

There are also quantitative versions, proved by Kalai [16] for balanced functions and by Mossel [23] for general functions (with tighter bounds obtained by Keller [18]). For simplicity we consider three alternatives and the impartial culture model. Then, the quantitative Arrow's theorem says that a reasonable pairwise comparison function f that is ε-far from every dictator (in the sense of normalized Hamming distance), must be such that the probability of Condorcet paradox is at least $\Omega(\varepsilon^3)$.

There is an analogous question about transitive dice: What are the methods for pairwise comparisons of k dice that always produce a linear order? In particular, we know that comparing two dice $\boldsymbol{a}$ and $\boldsymbol{b}$ by using the "beats" relation is not one of them.

We restrict ourselves to $k = 3$. Assume that we look at dice with n sides labeled with $[m]$, i.e., multisets of elements of $[m]$ of size n. Denote the set of such dice as $\mathcal{D}_{m,n}$. A pairwise comparison is an anti-symmetric function $f\colon (\mathcal{D}_{m,n} \times \mathcal{D}_{m,n}) \setminus \mathrm{diag}(\mathcal{D}_{m,n} \times \mathcal{D}_{m,n}) \to \{-1, 1\}$. We want to understand which pairwise comparison functions are transitive, i.e., there are no three distinct dice $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ such that $f(\boldsymbol{a}, \boldsymbol{b}) = f(\boldsymbol{b}, \boldsymbol{c}) = f(\boldsymbol{c}, \boldsymbol{a})$.

A little thought reveals that the answer is somewhat trivial. Let $\mathcal{O}$ be a linear order on $\mathcal{D}_{m,n}$. We think of $\mathcal{O}$ as an injective function $\mathcal{O}\colon \mathcal{D}_{m,n} \to \mathbb{R}$. If we define f as

$$f(\boldsymbol{a}, \boldsymbol{b}) = 1 \text{ if and only if } \mathcal{O}(\boldsymbol{a}) < \mathcal{O}(\boldsymbol{b}) \ ,$$

then f is easily seen to be transitive.

On the other hand, every transitive f must be of this form. To see this, consider a directed graph with vertex set $\mathcal{D}_{m,n}$ where there is an edge from $\boldsymbol{a}$ to $\boldsymbol{b}$ if and only if $f(\boldsymbol{a}, \boldsymbol{b}) = -1$. This graph is a tournament and transitivity of f means that it does not contain a directed triangle. But a triangle-free tournament does not contain a directed cycle and, therefore, induces a linear order on its ground set.

We can extend this reasoning to a quantitative result. It seems easiest to assume a model where a set of three dice is sampled u.a.r. from $\mathcal{D}_{m,n}$.

There is a result about tournaments due to Fox and Sudakov [8]. A tournament on n vertices is called *ε-far from transitive* if at least εn^2 of its edges must be reversed to obtain a transitive tournament.

Theorem 10 [8] *There exists $c > 0$ such that if a tournament on n vertices is ε-far from transitive, then it contains at least $c\varepsilon^2 n^3$ directed triangles.*

Theorem 10 can be restated as a quantitative Arrow-like statement for dice.

Corollary 3 *There exists $c > 0$ such that if a comparison function f on $\mathcal{D}_{m,n}$ with $m, n > 1$ is ε-far from transitive, then the probability that a random triple of dice is intransitive is at least $c\varepsilon^2$.*

Since [8] gives an example which is tight up to a constant factor, Corollary 3 is similarly tight. However, the obtained comparison function does not seem to correspond to any natural method of comparing dice.

Acknowledgements Open access funding provided by EPFL. We thank Timothy Gowers for helpful discussions of [32], Kathryn Mann for asking if there is an "Arrow's theorem" for dice, and the referee for a careful reading and helpful comments.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Missing calculations in the proof of Proposition 2

In this appendix we include, for ease of verification, several calculations that were omitted from the proof of Proposition 2.

Proof of Lemma 2

In the proof of the calculation lemma (Lemma 2) we performed the final calculation in (2.37) only to establish (2.20). Below we give analogous computations for (2.17)–(2.19) and (2.21)–(2.23).

Each of the three calculations establishing (2.17)–(2.19) proceeds in the same three steps: first, one of the conclusions (2.31)–(2.33) of Corollary 1 is applied; second, identities given in (2.35) and (2.36) are substituted for the integrals; third, the terms are rearranged.

For (2.17) we have, letting $D = \{(a_1, a_2, b_1, b_2) : a_1 > b_1 \wedge a_2 > b_2\}$,

$$\begin{aligned}\mathbb{P}[a_1 > b_1 \wedge a_2 > b_2 \mid \varepsilon_0] &= \left(1 + \frac{2}{n}\right) \iint_D f(a_1, a_2, b_1, b_2) \left(1 + \frac{2\alpha_2(a_1 + b_1)}{n} \right. \\ &\qquad \left. - \frac{a_1^2 + b_1^2 + a_1 a_2 + b_1 b_2}{n}\right) dab + o(n^{-1})\end{aligned}$$

$$= \left(1 + \frac{2}{n}\right)\left(\frac{1}{4} + \frac{\alpha_2 A}{n} - \frac{\alpha_2 A}{n} - \frac{B}{2n} - \frac{1-B}{2n} - \frac{A^2}{n} - \frac{A^2}{n}\right) + o(n^{-1})$$
$$= \frac{1}{4} - \frac{2A^2}{n} + o(n^{-1}).$$

For (2.18), letting $D = \{(a_1, b_1) : a_1 > b_1\}$

$$\mathbb{P}[a_1 > b_1 \mid \varepsilon_a] = \left(1 + \frac{1}{2n}\right)\iint_D f(a_1, b_1)\left(1 + \frac{\alpha_2 a_1}{n} - \frac{a_1^2}{2n}\right) dab + o(n^{-1})$$
$$= \left(1 + \frac{1}{2n}\right)\left(\frac{1}{2} + \frac{\alpha_2 A}{n} - \frac{B}{2n}\right) + o(n^{-1})$$
$$= \frac{1}{2} + \frac{1}{4n} + \frac{\alpha_2 A}{n} - \frac{B}{2n} + o(n^{-1}).$$

For (2.19), letting $D = \{(a_1, a_2, b_1, b_2 : a_1 > b_1 \wedge a_2 > b_2\}$,

$$\mathbb{P}[a_1 > b_1 \wedge a_2 > b_2 \mid \varepsilon_a]$$
$$= \left(1 + \frac{1}{n}\right)\iint_D f(a_1, a_2, b_1, b_2)\left(1 + \frac{2\alpha_2 a_1}{n} - \frac{a_1^2 + a_1 a_2}{n}\right) dab + o(n^{-1})$$
$$= \left(1 + \frac{1}{n}\right)\left(\frac{1}{4} + \frac{\alpha_2 A}{n} - \frac{B}{2n} - \frac{A^2}{n}\right) + o(n^{-1})$$
$$= \frac{1}{4} + \frac{1}{4n} + \frac{\alpha_2 A}{n} - \frac{B}{2n} - \frac{A^2}{n} + o(n^{-1}).$$

The calculations showing (2.21)–(2.23) employ Lemma 3 directly. Each of them applies one of (2.29)–(2.30) and uses the fact that both of those expressions can be approximated as

$$\iint_D f(a_1, a_2, b_1, b_2)\, dab + o(1).$$

More precisely, for (2.21) we take $D = \{(a_1, a_2, b_1, b_2) : a_1 > b_1 \wedge a_1 > b_2\}$ and get

$$\mathbb{P}[a_1 > b_1 \wedge a_1 > b_2 \mid \varepsilon_0] = \frac{\iint_D f(a_1, a_2, b_1, b_2)\varphi_{n-1}(-a_1)\varphi_{n-2}(-b_1 - b_2)\, dab}{\iint_{\mathbb{R}^4} f(a_1, a_2, b_1, b_2)\varphi_{n-1}(-a_1)\varphi_{n-2}(-b_1 - b_2)\, dab}$$
$$= \frac{\iint_D f(a_1, a_2, b_1, b_2)\, dab + o(1)}{\iint_{\mathbb{R}^4} f(a_1, a_2, b_1, b_2)\, dab + o(1)}$$
$$= \mathbb{P}[a_1 > b_1 \wedge a_1 > b_2] + o(1) = \frac{1}{3} + o(1).$$

Similarly, for (2.22), letting $D = \{(a_1, a_2, b_1, b_2) : a_1 > b_1 \wedge a_1 > b_2\}$,

$$\mathbb{P}[a_1 > b_1 \wedge a_1 > b_2 \mid \varepsilon_a] = \frac{\iint_D f(a_1, a_2, b_1, b_2)\varphi_{n-2}(-a_1 - a_2)\, dab}{\iint_{\mathbb{R}^4} f(a_1, a_2, b_1, b_2)\varphi_{n-2}(-a_1 - a_2)\, dab}$$

$$= \mathbb{P}[a_1 > b_1 \wedge a_1 > b_2] + o(1) = \frac{1}{3} + o(1),$$

and, for (2.23), letting $D = \{(a_1, a_2, b_1, c_1) : a_1 > b_1 \wedge a_1 > c_1\}$

$$\mathbb{P}\left[a_1 > b_1 \wedge a_1 > c_1 \mid \varepsilon_a \cap \varepsilon_b\right] = \frac{\iint_D f(a_1, a_2, b_1, c_1)\varphi_{n-2}(-a_1 - a_2)\varphi_{n-1}(-b_1)\, dabc}{\iint_{\mathbb{R}^4} f(a_1, a_2, b_1, c_1)\varphi_{n-2}(-a_1 - a_2)\varphi_{n-1}(-b_1)\, dabc}$$
$$= \mathbb{P}[a_1 > b_1 \wedge a_1 > c_1] + o(1) = \frac{1}{3} + o(1).$$

Proof of Lemma 3

In the proof of the integration lemma (Lemma 3), in (2.41) we included a detailed calculation only to establish (2.27). Below we give crucial steps of similar derivations for (2.28)–(2.30). In each of them: first, we substitute (2.40) for φ_j; second, we rearrange and absorb the error terms using (2.42) and (2.41).

For (2.28), we have

$$\begin{aligned}
&\iint_D f \cdot \varphi_{n-1}(-a)\, dab \\
&\quad = \iint_D f \cdot \left[\exp\left(-\frac{a^2}{2(n-1)}\right) \right. \\
&\qquad\qquad \left. \left(1 + \frac{\alpha_1}{n-1} + \frac{\alpha_2 a}{n-1} + O\left(\frac{\max(|a|, a^6)}{n^{3/2}}\right) + o(n^{-1})\right) \right] dab \\
&\quad = \iint_D f \cdot \left[1 + \frac{\alpha_1}{n} + \frac{\alpha_2 a}{n} - \frac{a^2}{2n}\right] dab + o(n^{-1}),
\end{aligned}$$

for (2.29),

$$\begin{aligned}
&\iint_D f \cdot \varphi_{n-2}(-a)\, dab \\
&\quad = \iint_D f \cdot \left[\exp\left(-\frac{a^2}{2(n-2)}\right) \right. \\
&\qquad\qquad \left. \left(1 + \frac{\alpha_1}{n-2} + \frac{\alpha_2 a}{n-2} + O\left(\frac{\max(|a|, a^6)}{n^{3/2}}\right) + o(n^{-1})\right) \right] dab \\
&\quad = \iint_D f \cdot \left[1 + \frac{\alpha_1}{n} + \frac{\alpha_2 a}{n} - \frac{a^2}{2n}\right] dab + o(n^{-1}),
\end{aligned}$$

and for (2.30),

$$\iint_D f \cdot \varphi_{n-2}(-a)\varphi_{n-1}(-b)\, dabc$$

$$= \iint_D f \cdot \left[\exp\left(-\frac{a^2}{2(n-2)} \right) \left(1 + \frac{\alpha_1}{n-2} + \frac{\alpha_2 a}{n-2} + O\left(\frac{\max(|a|, a^6)}{n^{3/2}} \right) + o(n^{-1}) \right) \right]$$
$$\cdot \left[\exp\left(-\frac{b^2}{2(n-1)} \right) \right.$$
$$\left. \left(1 + \frac{\alpha_1}{n-1} + \frac{\alpha_2 b}{n-1} + O\left(\frac{\max(|b|, b^6)}{n^{3/2}} \right) + o(n^{-1}) \right) \right] dabc$$
$$= \iint_D f \cdot \left[1 + \frac{2\alpha_1}{n} + \frac{\alpha_2(a+b)}{n} - \frac{a^2+b^2}{2n} \right] dabc + o(n^{-1}).$$

References

1. Arndt, J., Haenel, C.: Pi—Unleashed. Translated from the 1998 German Original by Catriona Lischka and David Lischka, 2nd edn. Springer-Verlag, Berlin (2001). https://doi.org/10.1007/978-3-642-56735-3
2. Arrow, K.J.: A difficulty in the concept of social welfare. J. Political Econ. **58**(4), 328–346 (1950)
3. Arrow, K.J.: Social Choice and Individual Values. No. 12 in Cowles Foundation Monographs. Wiley, Hoboken (1963)
4. Bentkus, V.: A Lyapunov type bound in $\mathbb{R}^d$. Theory Prob. Appl. **49**(2), 311–323 (2005). https://doi.org/10.1137/S0040585X97981123
5. Bhattacharya, R.N., Rao, R.R.: Normal Approximation and Asymptotic Expansions. Society for Industrial and Applied Mathematics, Philadelphia (2010). https://doi.org/10.1137/1.9780898719895
6. Breuer, P., Major, P.: Central limit theorems for non-linear functionals of Gaussian fields. J. Multivariate Anal. **13**(3), 425–441 (1983). https://doi.org/10.1016/0047-259X(83)90019-2
7. Conrey, B., Gabbard, J., Grant, K., Liu, A., Morrison, K.E.: Intransitive dice. Math. Mag. **89**(2), 133–143 (2016)
8. Fox, J., Sudakov, B.: Unavoidable patterns. J. Comb. Theory Ser. A **115**(8), 1561–1569 (2008). https://doi.org/10.1016/j.jcta.2008.04.003
9. Friedgut, E., Kalai, G., Keller, N., Nisan, N.: A quantitative version of the Gibbard–Satterthwaite theorem for three alternatives. SIAM J. Comput. **40**(3), 934–952 (2011). https://doi.org/10.1137/090756740
10. Gardner, M.: Mathematical games: the paradox of the nontransitive dice and the elusive principle of indifference. Sci. Am. **223**, 110–114 (1970)
11. Garman, M.B., Kamien, M.I.: The paradox of voting: probability calculations. Behav. Sci. **13**(4), 306–316 (1968). https://doi.org/10.1002/bs.3830130405
12. Gehrlein, W.V.: Condorcet's paradox and the likelihood of its occurrence: different perspectives on balanced preferences. Theory Decis. **52**(2), 171–199 (2002). https://doi.org/10.1023/A:1015551010381
13. Guilbaud, G.T.: Les théories de l'intérêt général et le problème logique de l'agrégation. Economie appliquée **V**(4), 501–551 (1952)
14. Isaksson, M., Kindler, G., Mossel, E.: The geometry of manipulation—a quantitative proof of the Gibbard Satterthwaite theorem. Combinatorica **32**(2), 221–250 (2012)
15. Isaksson, M., Mossel, E.: New maximally stable Gaussian partitions with discrete applications. Isr. J. Math. **189**, 347–396 (2012)
16. Kalai, G.: A Fourier-theoretic perspective on the Condorcet paradox and Arrow's theorem. Adv. Appl. Math. **29**(3), 412–426 (2002). https://doi.org/10.1016/S0196-8858(02)00023-4
17. Kalai, G.: Noise sensitivity and chaos in social choice theory. In: Katona, G.O.H., Schrijver, A., Szonyi, T., Sagi, G. (eds.) Fete of Combinatorics and Computer Science, pp. 173–212. Springer, Berlin (2010)
18. Keller, N.: A tight quantitative version of Arrow's impossibility theorem. J. Eur. Math. Soc. **14**(5), 1331–1355 (2012)
19. Lowe, J.: Bill Gates Speaks: Insight from the World's Greatest Entrepreneur. Wiley, Hoboken (2001)
20. Marinucci, D., Wigman, I.: On the area of excursion sets of spherical Gaussian eigenfunctions. J. Math. Phys. **52**(093301), 1–21 (2011). https://doi.org/10.1063/1.3624746
21. May, R.M.: Some mathematical remarks on the paradox of voting. Behav. Sci. **16**(2), 143–151 (1971). https://doi.org/10.1002/bs.3830160204

22. Mossel, E.: Gaussian bounds for noise correlation of functions. Geom. Funct. Anal. **19**(6), 1713–1756 (2010). https://doi.org/10.1007/s00039-010-0047-x
23. Mossel, E.: A quantitative Arrow theorem. Probab. Theory Relat. Fields **154**(1), 49–88 (2012). https://doi.org/10.1007/s00440-011-0362-7
24. Mossel, E., O'Donnell, R., Oleszkiewicz, K.: Noise stability of functions with low influences: invariance and optimality. Ann. Math. **171**(1), 295–341 (2010). https://doi.org/10.4007/annals.2010.171.295
25. Mossel, E., Rácz, M.Z.: A quantitative Gibbard–Satterthwaite theorem without neutrality. Combinatorica **35**(3), 317–387 (2015)
26. Niemi, R.G., Weisberg, H.F.: A mathematical solution for the probability of the paradox of voting. Behav. Sci. **13**(4), 317–323 (1968). https://doi.org/10.1002/bs.3830130406
27. Nourdin, I.: Selected Aspects of Fractional Brownian Motion, 1st edn. Springer, Milano (2012). https://doi.org/10.1007/978-88-470-2823-4
28. Nourdin, I., Peccati, G.: Normal Approximations with Malliavin Calculus, Cambridge Tracts in Mathematics, vol. 192. Cambridge University Press, Cambridge (2012)
29. O'Donnell, R.: Analysis of Boolean Functions. Cambridge University Press, Cambridge (2014)
30. Petrov, V.V.: Sums of Independent Random Variables. Translated from the Russian by A, p. 82. A. Brown, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band. Springer-Verlag, New York (1975)
31. Polymath, D.: Intransitive dice VII—aiming for further results (2017). https://gowers.wordpress.com/2017/08/12/intransitive-dice-vii-aiming-for-further-results/. Accessed 1 Apr 2020
32. Polymath, D.: Polymath 13—a success! (2017). https://polymathprojects.org/2017/08/22/polymath-13-a-success/. Accessed 1 Apr 2020
33. Ramanujan, S.: On question 330 of Professor Sanjana [J. Indian Math. Soc. **4** (1912), 59–61]. In: Collected papers of Srinivasa Ramanujan, pp. 15–17. AMS Chelsea Publ., Providence, RI (2000)
34. Spitzer, F.: Principles of Random Walk, 2nd edn. Springer-Verlag, New York (1976). https://doi.org/10.1007/978-1-4757-4229-9

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Jan Hązła[1] · Elchanan Mossel[2] · Nathan Ross[3] · Guangqu Zheng[4]

Elchanan Mossel
elmos@mit.edu

Nathan Ross
nathan.ross@unimelb.edu.au

Guangqu Zheng
zhengguangqu@gmail.com

1 EPFL, Lausanne, Switzerland

2 Massachusetts Institute of Technology, Cambridge, USA

3 University of Melbourne, Parkville, Australia

4 University of Kansas, Lawrence, USA