Combinatorics, Probability and Computing (2021), **30**, pp. 398–411
doi:10.1017/S0963548320000474

CAMBRIDGE
UNIVERSITY PRESS

ARTICLE

A simplified disproof of Beck's three permutations conjecture and an application to root-mean-squared discrepancy

Cole Franks[†]

Department of Mathematics, Rutgers University, Piscataway, NJ 08854, USA
Email: franks@mit.edu

(Received 2 November 2018; accepted 24 September 2019; revised 6 August 2020; first published online 26 October 2020)

Abstract

A k-permutation family on n vertices is a set-system consisting of the intervals of k permutations of the integers 1 to n. The discrepancy of a set-system is the minimum over all red–blue vertex colourings of the maximum difference between the number of red and blue vertices in any set in the system. In 2011, Newman and Nikolov disproved a conjecture of Beck that the discrepancy of any 3-permutation family is at most a constant independent of n. Here we give a simpler proof that Newman and Nikolov's sequence of 3-permutation families has discrepancy $\Omega(\log n)$. We also exhibit a sequence of 6-permutation families with root-mean-squared discrepancy $\Omega(\sqrt{\log n})$; that is, in any red–blue vertex colouring, the square root of the expected squared difference between the number of red and blue vertices in an interval of the system is $\Omega(\sqrt{\log n})$.

2020 MSC Codes: Primary: 68R01; Secondary: 05D99

1. Introduction

The discrepancy of a set-system is the extent to which the sets in a set-system can be simultaneously split into two equal parts, or two-coloured in a balanced way. Let $\mathcal{A}$ be a collection (possibly with multiplicity) of subsets of a finite set Ω. The discrepancy of a two-colouring $\chi : \Omega \to \{\pm 1\}$ of the set-system $(\Omega, \mathcal{A})$ is the maximum imbalance in colour over all sets S in $\mathcal{A}$. The discrepancy of $(\Omega, \mathcal{A})$ is the minimum discrepancy of any two-colouring of Ω. Formally,

$$\operatorname{disc}_\infty(\Omega, \mathcal{A}) := \min_{\chi : \Omega \to \{+1,-1\}} \operatorname{disc}_\infty(\chi, \mathcal{A}), \tag{1.1}$$

where

$$\operatorname{disc}_\infty(\chi, \mathcal{A}) = \max_{S \in \mathcal{A}} |\chi(S)| \quad \text{and} \quad \chi(S) = \sum_{x \in S} \chi(x).$$

A central goal of the study of discrepancy is to bound the discrepancy of set-systems with restrictions or additional structure. Here we will be concerned with set-systems constructed from permutations. A permutation $\sigma : \Omega \to \Omega$ from a set Ω with a total ordering $\leqslant$ to itself determines the set-system $(\Omega, \mathcal{A}_\sigma)$, where

[†]Supported in part by Simons Foundation award 332622.

$$\mathcal{A}_\sigma = \{\{i\colon \sigma(i) \leqslant \sigma(j)\}\colon j \in \Omega\} \cup \{\emptyset\}.$$

For example, if [3] inherits the usual ordering on natural numbers and $e\colon [3] \to [3]$ is the identity permutation, then $\mathcal{A}_e = \{\emptyset, \{1\}, \{1,2\}, \{1,2,3\}\}$. Equivalently, $\mathcal{A}_\sigma$ is a maximal chain in the poset $2^{[n]}$ ordered by inclusion. If $P = \{\sigma_1, \ldots, \sigma_k\}$ is a set of permutations of Ω, let $\mathcal{A}^P = \mathcal{A}_{\sigma_1} + \cdots + \mathcal{A}_{\sigma_k}$, where $+$ denotes multiset sum (union with multiplicity). Then we say $(\Omega, \mathcal{A}^P)$ is a *k-permutation family*.

By Dilworth's theorem, the maximal discrepancy of a k-permutation family is the same as the maximal discrepancy of a set-system of *width* k, that is, a set-system that contains no antichain of cardinality more than k.

It is easy to see that a 1-permutation family has discrepancy at most 1, and the same is true for 2-permutation families [9]. Beck conjectured that the discrepancy of a 3-permutation family is $O(1)$. More generally, Spencer, Srinivasan and Tetali [10] conjectured that the discrepancy of a k-permutation family is $O(\sqrt{k})$. Both conjectures were recently disproved by Newman and Nikolov [7]. They showed the following.

Theorem 1.1 ([7]). *There is a sequence of* 3*-permutation families on n vertices with discrepancy* $\Omega(\log n)$.

The same authors together with Neiman [6] showed that the above lower bound implies that a natural class of rounding schemes for the Gilmore–Gomory linear programming relaxation of bin-packing (such as the scheme used in the Kamarkar–Karp algorithm) incur logarithmic error.

Spencer, Srinivasan and Tetali [10] proved an upper bound that matches the lower bound of Newman and Nikolov for $k = 3$.

Theorem 1.2 ([10]). *The discrepancy of a k-permutation family on n vertices is* $O(\sqrt{k}\log n)$.

For very large k, in particular $k \geqslant n$, Spencer, Srinivasan and Tetali [10] showed that the discrepancy is of the order of $\min\{n, \sqrt{n \log(2k/n)}\}$. For $3 < k = o(n)$, however, the best known lower bound is the trivial $\Omega(\max \sqrt{k}, \log n)$, leaving open the tightness of Theorem 1.2. Even the following weaker conjecture is open.

Conjecture 1.3 (Michael Saks). *There is a function* $f \in \omega(1)$ *such that there is a family* $M_{n,k}$ *of k-permutation families on n vertices satisfying*

$$\operatorname{disc}_\infty(M_{n,k}) \geqslant f(k) \log n$$

for all $k, n \in \mathbb{N}$ *with* $k \leqslant n$.

In this paper we present a new analysis of the counterexample due to Newman and Nikolov. We replace their case analysis with a simple argument using norms of matrices, albeit achieving a worse constant $((4\sqrt{6})^{-1}\log_3 n$ versus their $3^{-1}\log_3 n)$. Our analysis generalizes well to larger-permutation families, and can hopefully be extended to handle Conjecture 1.3. Our analysis also yields a new result for the root-mean-squared discrepancy, defined as

$$\operatorname{disc}_2(\Omega, \mathcal{A}) = \min_{\chi\colon \Omega \to \{\pm 1\}} \operatorname{disc}_2(\mathcal{A}, \chi),$$

where

$$\operatorname{disc}_2(\mathcal{A}, \chi) = \sqrt{\frac{1}{|\mathcal{A}|} \sum_{S \in \mathcal{A}} |\chi(S)|^2}.$$

https://doi.org/10.1017/S0963548320000474 Published online by Cambridge University Press

Define the *hereditary root-mean-squared discrepancy* by

$$\operatorname{herdisc}_2(\Omega, \mathcal{A}) = \max_{\Gamma \subset \Omega} \operatorname{disc}_2(\Gamma, \mathcal{A}|_\Gamma).$$

Theorem 1.4. *There is a sequence of* 6*-permutation families on n vertices with root-mean-squared discrepancy* $\Omega(\sqrt{\log n})$.

For $k = 6$, Theorem 1.4 matches the upper bound of $\sqrt{k \log n}$ for the root-mean-squared discrepancy implicitly proved in [10]. Further, in [3] it is shown that a certain efficiently computable quantity approximates the hereditary root-mean-squared discrepancy up to a factor of $\sqrt{\log n}$. As communicated to the author by Aleksandar Nikolov, this quantity is constant for families of constantly many permutations. Thus Theorem 1.4 shows that the $\sqrt{\log n}$ gap between $\operatorname{herdisc}_2(\Omega, \mathcal{A})$ and the quantity defined in [3] is best possible.

Remark 1.1 (odd discrepancy). The proofs of our lower bounds, Theorem 1.1 and Theorem 1.4, only use that the assignment $\chi : \Omega \to \{\pm 1\}$ assigns *odd integers* to Ω. As such, the lower bounds still hold for the relaxed notion of discrepancy in which we allow $\chi : \Omega \to 2\mathbb{Z} - 1$.

2. The set-system of Newman and Nikolov

Our proof of Theorem 1.1 uses the same set-system as Newman and Nikolov. For completeness, we define and slightly generalize the system here. The vertices of the system will be r-ary strings, or elements of $[r]^d$. For Newman and Nikolov's set-system, $r = 3$. We first set our notation for referring to strings.

Definition 2.1 (string notation).

- Bold letters, *e.g.* $\boldsymbol{a}$, denote strings in $[r]^d$ for some $d \geqslant 0$. Here $[r]^0$ denotes the set containing only the empty string $\boldsymbol{\varepsilon}$.
- If $\boldsymbol{a} = a_1 \ldots a_d \in [r]^d$ is a string, for $0 \leqslant k \leqslant d$ let $\boldsymbol{a}[k]$ denote the string $a_1 \ldots a_k$, with $\boldsymbol{a}[0] := \boldsymbol{\varepsilon}$.
- If $\boldsymbol{a}$ is a string, $|\boldsymbol{a}|$ denotes the length of $\boldsymbol{a}$.
- If $\boldsymbol{a}$ and $\boldsymbol{b}$ are strings, their concatenation in $[r]^{|\boldsymbol{a}|+|\boldsymbol{b}|}$ is denoted $\boldsymbol{ab}$.
- If $j \in [r]$, then $\bar{j}$ denotes the all-j string of length d, for example

$$\bar{3} := \underbrace{33\ldots3}_{d}.$$

- τ denotes the permutation of $[r]$ given by $\tau(i) = r - i + 1$, the permutation reversing the ordering on $[r]$.

We may now define the set-system of Newman and Nikolov.

Definition 2.2 (set-system $([r]^d, \mathcal{A}_P)$). Let $<$ be the lexicographical ordering on $[r]^d$. Given a permutation σ of $[r]$, we define a permutation $\boldsymbol{\sigma}$ of $[r]^d$ by acting digitwise by σ. Namely, $\sigma(\boldsymbol{a}) := \sigma(a_1)\sigma(a_2)\ldots\sigma(a_d)$. For any subset $P \subset S_r$ of permutations of $[r]$, define the permutation family $\mathcal{A}_P$ by

$$\mathcal{A}_P = \mathcal{A}^{\{\boldsymbol{\sigma} : \sigma \in P\}}.$$

https://doi.org/10.1017/S0963548320000474 Published online by Cambridge University Press

Namely, the edges of $\mathcal{A}_P$ are $\emptyset$ and the sets $\leqslant_\sigma \boldsymbol{a}$ defined by

$$\leqslant_\sigma \boldsymbol{a} := \{\boldsymbol{b} \in [r]^{|\boldsymbol{a}|} : \boldsymbol{\sigma}(\boldsymbol{b}) \leqslant \boldsymbol{\sigma}(\boldsymbol{a})\}$$

as σ ranges over P and $\boldsymbol{a}$ over $[r]^d$. Note that for each $\sigma \in P$ and $\boldsymbol{a} \in [r]^d$, $\mathcal{A}_P$ also contains the edges defined not to include $\boldsymbol{a}$:

$$<_\sigma \boldsymbol{a} := \{\boldsymbol{b} \in [r]^{|\boldsymbol{a}|} : \boldsymbol{\sigma}(\boldsymbol{b}) < \boldsymbol{\sigma}(\boldsymbol{a})\}.$$

Definition 2.3 (set-system of Newman and Nikolov). The system of Newman and Nikolov is $([3]^d, \mathcal{A}_C)$ with $C = \{e, (1, 2, 3), (1, 3, 2)\}$. That is, C is the cyclic permutations of 3.

We first bound the discrepancy of the 6-permutation family $([3]^d, \mathcal{A}_{S_3})$. In fact, we bound the smaller *odd* discrepancy of this family.

Theorem 2.1 (discrepancy lower bound). *If $r \geqslant 3$ is odd, then*

$$\operatorname{disc}_\infty([r]^d, \mathcal{A}_{S_r}) \geqslant \frac{d}{2\sqrt{6}}.$$

Theorem 2.1 is proved in the next section, Section 2.1. We bound the discrepancy of $\mathcal{A}_C$ in terms of the discrepancy of $\mathcal{A}_{S_3}$. Theorem 1.1 follows immediately from Proposition 2.1 and Theorem 2.1 for $r = 3$.

Proposition 2.1 ([7]).

$$\operatorname{disc}_\infty([3]^d, \mathcal{A}_C) \geqslant \frac{1}{2} \operatorname{disc}_\infty([3]^d, \mathcal{A}_{S_3}).$$

Proof. The key observation is that $[3]^d$ is in reverse order under the action of σ and $\tau \circ \sigma$, so for each $\sigma, \boldsymbol{a}$, the edges $\leqslant_\sigma \boldsymbol{a}$ and $<_{\tau\circ\sigma} \boldsymbol{a}$ partition $[3]^d$.

Let $\chi : [3]^d \to 2\mathbb{Z} - 1$ be an assignment of minimal discrepancy K to $([3]^d, \mathcal{A}_C)$. Let $\sigma \in S_3, \boldsymbol{a} \in [r]^d$ be arbitrary. It suffices to show $|\chi(\leqslant_\sigma \boldsymbol{a})| \leqslant 2K$. If $\sigma \in C$, then $|\chi(\leqslant_\sigma \boldsymbol{a})| \leqslant K$ and so the claim is trivial. If not, then $\tau \circ \sigma \in C$. By the above reasoning,

$$|\chi(\leqslant_\sigma \boldsymbol{a})| = |\chi([3]^d) - \chi(<_{\tau\circ\sigma} \boldsymbol{a})| \leqslant 2K. \qquad \square$$

We recall a few facts about the set-system of Newman and Nikolov. As observed in [7], the quantity $\chi(<_\sigma \boldsymbol{a})$ behaves additively under concatenation of strings. That is, if $\boldsymbol{a} = \boldsymbol{bc}$, then there is a natural way to define colourings $\chi^{\boldsymbol{\varepsilon}}$ and $\chi^{\boldsymbol{b}}$ of $[r]^{|\boldsymbol{b}|}$ and $[r]^{|\boldsymbol{c}|}$, respectively, such that $\chi(<_\sigma \boldsymbol{a}) = \chi^{\boldsymbol{\varepsilon}}(<_\sigma \boldsymbol{b}) + \chi^{\boldsymbol{b}}(<_\sigma \boldsymbol{c})$. Further, this holds even if χ is a colouring of $[r]^d$ with any odd integers (rather than just ± 1).

Definition 2.4 (extension of colourings). We extend each assignment $\chi : [r]^d \to 2\mathbb{Z} - 1$ to

$$\chi : [r]^0 \cup [r]^1 \cup \cdots \cup [r]^d \to 2\mathbb{Z} - 1$$

by defining

$$\chi(\boldsymbol{a}) = \sum_{|\boldsymbol{b}| = d - |\boldsymbol{a}|} \chi(\boldsymbol{ab}) \quad \text{for } |\boldsymbol{a}| \leqslant d.$$

Crucially, the entries of χ are indeed odd. Observe that

$$\chi(\boldsymbol{a}) = \sum_{i \in [r]} \chi(\boldsymbol{a}i) \tag{2.1}$$

https://doi.org/10.1017/S0963548320000474 Published online by Cambridge University Press

for $|\boldsymbol{a}| < d$. For $|\boldsymbol{a}| \leqslant d$, define

$$\chi^{\boldsymbol{a}}\colon [r]^0 \cup [r]^1 \cup \cdots \cup [r]^{d-|\boldsymbol{a}|}$$

by $\chi^{\boldsymbol{a}}(\boldsymbol{b}) = \chi(\boldsymbol{ab})$ for $|\boldsymbol{b}| \leqslant d - |\boldsymbol{a}|$. In particular, $\chi^{\boldsymbol{\varepsilon}}$ and χ are equal as functions on $[r]^0 \cup [r]^1 \cup \cdots \cup [r]^d$.

Proposition 2.2 (additivity of discrepancy). *For any assignment $\chi\colon [r]^d \to 2\mathbb{Z} - 1$ and $\boldsymbol{a} = \boldsymbol{bc}$ with $|\boldsymbol{a}| \leqslant d$, we have*

$$\chi(<_\sigma \boldsymbol{a}) = \chi^{\boldsymbol{\varepsilon}}(<_\sigma \boldsymbol{b}) + \chi^{\boldsymbol{b}}(<_\sigma \boldsymbol{c}) \tag{2.2}$$

Proof. If $|\boldsymbol{b}'| = |\boldsymbol{b}|$ and $|\boldsymbol{c}| = |\boldsymbol{c}'|$, then $\boldsymbol{b}'\boldsymbol{c}'$ is in $\leqslant_\sigma \boldsymbol{bc}$ if and only if $\sigma \cdot \boldsymbol{b}' < \sigma \cdot \boldsymbol{b}$ or $\boldsymbol{b}' = \boldsymbol{b}$ and $\sigma \cdot \boldsymbol{c}' < \sigma \cdot \boldsymbol{c}$. Thus

$$\chi(\leqslant_\sigma \boldsymbol{bc}) = \sum_{\sigma\cdot\boldsymbol{b}'<\sigma\cdot\boldsymbol{b}} \sum_{|\boldsymbol{c}'|=|\boldsymbol{c}|} \chi(\boldsymbol{b}'\boldsymbol{c}') + \sum_{\sigma\cdot\boldsymbol{c}'<\sigma\cdot\boldsymbol{c}} \chi(\boldsymbol{bc}').$$

The right-hand side is precisely $\chi^{\boldsymbol{\varepsilon}}(<_\sigma \boldsymbol{b}) + \chi^{\boldsymbol{b}}(<_\sigma \boldsymbol{c})$. □

2.1 Proof of the lower bound

In this section we prove Theorem 2.1, the lower bound on $\operatorname{disc}_\infty([r]^d, \mathcal{A}_{S_r})$. We now describe the plan of the proof and use it to motivate several definitions. The proof appears afterwards at the end of the section.

To show the discrepancy $\operatorname{disc}_\infty([r]^d, \mathcal{A}_{S_r})$ is at least K, it is enough to show (because, in particular, ± 1 are odd) that given an assignment $\chi\colon [r]^d \to 2\mathbb{Z} - 1$, we can choose $\sigma \in S_r$ and $\boldsymbol{a} \in [r]^d$ so that $|\chi(<_\sigma \boldsymbol{a})|$ is at least K.

We do this in two steps. First, define a vector $M_\chi(\boldsymbol{a})$ depending on χ and the choice of $\boldsymbol{a}$ and an appropriate norm $\|\cdot\|$ such that if $\|M_\chi(\boldsymbol{a})\| \geqslant K$, then there exists σ with $|\chi(<_\sigma \boldsymbol{a})| \geqslant K$. Next, we choose $\boldsymbol{a}$ to maximize $\|M_\chi(\boldsymbol{a})\|$. The correct object M_χ turns out to be an $r \times r$ matrix-valued function of $\boldsymbol{a}$, and rather than a norm we use a seminorm denoted by $\|\cdot\|_{S_r}$. We will define the two such that, for any $\boldsymbol{a} \in [r]^d$,

$$\max_\sigma |\chi(<_\sigma \boldsymbol{a})| = \|M_\chi(\boldsymbol{a})\|_{S_r}.$$

Definition 2.5 (seminorm $\|\cdot\|_{S_r}$). For $M \in \operatorname{Mat}_{r\times r}(\mathbb{R})$ and $\sigma \in S_r$, define

$$\sigma \cdot M := \sum_{i,j\in[r],\, \sigma(i)>\sigma(j)} M_{i,j}.$$

Now let $\|M\|_{S_r} = \max_{\sigma\in S_r} |\sigma \cdot M|$.

Remark 2.1. This seminorm is well studied; if M is the 0, 1 adjacency matrix of a directed graph G, then $\|M\|_{S_r}$ is the maximum size of an acyclic subgraph of G. In [2] it is shown that, assuming the unique games conjecture, $\|M\|_{S_r}$ is **NP**-hard to approximate even for M antisymmetric.

It remains to define the matrix M. We will define M such that

$$\chi(<_\sigma \boldsymbol{a}) = \sigma \cdot M_\chi(\boldsymbol{a}). \tag{2.3}$$

Recall how we extended χ in Definition 2.4. If we define M_χ to be an *additive* function on $[r]^0 \cup [r]^1 \cup \cdots \cup [r]^d$, *i.e.* one satisfying

$$M_\chi(\boldsymbol{ab}) = M_{\chi^{\boldsymbol{\varepsilon}}}(\boldsymbol{a}) + M_{\chi^{\boldsymbol{a}}}(\boldsymbol{b}), \tag{2.4}$$

https://doi.org/10.1017/S0963548320000474 Published online by Cambridge University Press

then by linearity of $\sigma \cdot M$ in M we only need to check that (2.3) holds for $d = 1$. This motivates our definition of M_χ.

Definition 2.6 (matrix $M_\chi(\boldsymbol{a})$). Let $\chi : [r]^d \to 2\mathbb{Z} - 1$. For $d = 0$, define $M_\chi(\boldsymbol{\varepsilon}) = 0$. For $a \in [r]$, *i.e.* $d = 1$, define $M_\chi(a)$ to be the $r \times r$ matrix with only the ath row non-zero, and the entries of this row given by $\chi(1), \chi(2) \ldots, \chi(r)$. Equivalently,

$$M_\chi(a)_{i,j} = \delta_{i,a}\chi(j) \quad \text{for } a, i, j \in [r]. \tag{2.5}$$

For $d > 1$, define

$$M_\chi(\boldsymbol{a}) = \sum_{k=1}^{|\boldsymbol{a}|} M_{\chi^{\boldsymbol{a}[k-1]}}(a_k). \tag{2.6}$$

For example, suppose $d = 2$ and $\chi(11) = \chi^1(1) = 1$, $\chi(12) = \chi^1(2) = -1$, $\chi(13) = \chi^1(3) = 1$ so that $\chi(1) = \chi^\varepsilon(1) = 1 - 1 + 1 = 1$, and suppose also that $\chi(2) = \chi^\varepsilon(2) = 3$ and $\chi(3) = \chi^\varepsilon(3) = -3$. Then

$$M_\chi(12) = M_{\chi^\varepsilon}(1) + M_{\chi^1}(2) = \begin{bmatrix} 1 & 3 & -3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 1 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

We now prove that the matrix and seminorm we have defined have the promised property.

Proposition 2.3. *For all $\chi : [r]^d \to 2\mathbb{Z} - 1$, (2.3) holds, and hence*

$$\max_\sigma |\chi(<_\sigma \boldsymbol{a})| = \|M_\chi(\boldsymbol{a})\|_{S_r}.$$

Proof of Proposition 2.3. First, we claim that M_χ is additive, that is, (2.4) holds. Indeed,

$$M_\chi(\boldsymbol{ab}) = \sum_{k=1}^{|\boldsymbol{a}|} M_{\chi^{\boldsymbol{a}[k-1]}}(a_k) + \sum_{k=1}^{|\boldsymbol{b}|} M_{\chi^{\boldsymbol{a}(\boldsymbol{b}[k-1])}}(b_k) = \sum_{k=1}^{|\boldsymbol{a}|} M_{\chi^{\boldsymbol{a}[k-1]}}(a_k) + \sum_{k=1}^{|\boldsymbol{b}|} M_{(\chi^{\boldsymbol{a}})^{\boldsymbol{b}[k-1]}}(b_k),$$

which is exactly $M_{\chi^\varepsilon}(\boldsymbol{a}) + M_{\chi^{\boldsymbol{a}}}(\boldsymbol{b})$. By additivity, it is enough to prove that (2.3) holds for $d = 1$. This is a straightforward calculation. By (2.5), for $a \in [r]$,

$$\sigma \cdot M_\chi(a) = \sum_{i,j \in [r],\, \sigma(i) > \sigma(j)} \delta_{i,a}\chi(j) = \sum_{j \in [r]:\ \sigma(j) < \sigma(a)} \chi(j).$$

The right-hand side is exactly $\chi(<_\sigma a)$. □

Now that we have Proposition 2.3, it remains to bound $\min_\chi \max_{\boldsymbol{a}} \|M_\chi(\boldsymbol{a})\|_{S_r}$ below. This quantity is at least the value of the following d-round game played between a 'minimizer' and a 'maximizer'.

Definition 2.7 (seminorm unbalancing game). The states of the seminorm balancing game are $r \times r$ integer matrices M. The matrix M is updated in each round as follows.

(1) The minimizer chooses a row vector v in $(2\mathbb{Z} - 1)^r$, *i.e.* a list of r odd numbers.
(2) The maximizer chooses a number $i \in [r]$ and adds v to the ith row of M.

The value for the maximizer is the value of $\|M\|_{S_r}$ at the end of the game.

https://doi.org/10.1017/S0963548320000474 Published online by Cambridge University Press

We now discuss why the value of this game is a lower bound on $\min_\chi \max_{\boldsymbol{a}} \|M_\chi(\boldsymbol{a})\|_{S_r}$. The sequence of moves made by the maximizer can be viewed as a string $\boldsymbol{a} \in [r]^d$, and we claim that the assignment $\chi : [r]^d \to 2\mathbb{Z} - 1$ determines a strategy such that the value for the maximizer is exactly $\|M_\chi(\boldsymbol{a})\|_{S_r}$.

Here is how a colouring $\chi : [r]^d \to 2\mathbb{Z} - 1$ determines a strategy for the minimizer. If the maximizer chose rows $\boldsymbol{a} = a_1, \ldots, a_{k-1}$ in rounds $1, \ldots, k-1$, the minimizer chooses the vector $v = \chi(\boldsymbol{a}1), \ldots, \chi(\boldsymbol{a}r)$ in round k, where χ on $[r]^{k+1}$ is determined by χ on $[r]^d$ as in Definition 2.4. If the minimizer plays this strategy and the maximizer plays $\boldsymbol{a} \in [r]^d$, the matrix after the kth round will be $M_\chi(\boldsymbol{a}[k])$, because $M_{\chi^{\boldsymbol{a}[k-1]}}(a_k)$ has v in the a_kth row and zeros elsewhere. If the minimizer is constrained to choose w, v in the $(k-1)$th and kth rounds, respectively, such that $\sum_{i=1}^r v_i = w_{a_{k-1}}$, then by (2.2) the strategy of the minimizer is determined by some colouring χ as above. However, we show that the value of the game is $\Omega(d)$ even without this constraint on the minimizer.

To compute a lower bound on the value of the game, we first bound the seminorm below by a simpler quantity. Recall that $\|M\|_F$ denotes the Frobenius norm of the matrix M, *i.e.* it is the square root of the sum of squares of its entries.

Lemma 2.4. *For $\sigma \in S_r$ chosen uniformly at random,*

$$\|M\|_{S_r} \geqslant \sqrt{\mathbb{E}_\sigma(\sigma \cdot M)^2} \geqslant \frac{1}{2\sqrt{6}} \|M - M^T\|_F.$$

Proof of Lemma 2.4. The first inequality is immediate. Let T be the matrix with ones strictly below the main diagonal and zeros elsewhere. For the second inequality, we use the identity

$$\mathbb{E}_\sigma(\sigma \cdot M)^2 = \frac{1}{4}(\operatorname{Tr} T(M + M^T))^2 + \frac{1}{4}\mathbb{E}_\sigma(\sigma \cdot (M - M^T))^2. \tag{2.7}$$

Equation (2.7) follows because the expectation of the square of a random variable is its mean squared plus its variance, and $\mathbb{E}_\sigma \sigma \cdot M = \frac{1}{2} \operatorname{Tr} T(M + M^T)$. The second term is the variance because $M = \frac{1}{2}(M + M^T) + \frac{1}{2}(M - M^T)$, and for any $\tau \in S_r$ we have $\tau \cdot \frac{1}{2}(M + M^T) = \mathbb{E}_\sigma \sigma \cdot M$.

Set $A = M - M^T$. In particular, A is antisymmetric. Write

$$\begin{aligned}
\mathbb{E}_\sigma(\sigma \cdot A)^2 &= \sum_{i \neq j, k \neq \ell} A_{i,j} A_{k,l} \mathbb{E}[\mathbf{1}_{\sigma(i) > \sigma(j)} \mathbf{1}_{\sigma(k) > \sigma(\ell)}] \\
&= \frac{1}{4} \sum_{|\{i,j,k,\ell\}|=4} A_{i,j} A_{k,l} + \frac{1}{3} \sum_{|\{i,j,\ell\}|=3} 2A_{i,j} A_{i,\ell} \\
&\quad + \frac{1}{6} \sum_{|\{i,j,\ell\}|=3} 2A_{i,j} A_{j,\ell} + \frac{1}{2} \sum_{|\{i,j\}|=2} A_{i,j} A_{i,j}.
\end{aligned}$$

This expression is obtained by computing $\mathbb{E}[\mathbf{1}_{\sigma(i) > \sigma(j)} \mathbf{1}_{\sigma(k) > \sigma(\ell)}]$ in each of the cases and using antisymmetry of A.

- If $|\{i, j, k, \ell\}| = 4$, then $\mathbb{E}[\mathbf{1}_{\sigma(i) > \sigma(j)} \mathbf{1}_{\sigma(k) > \sigma(\ell)}] = 1/4$. This gives us the first term.
- If $|\{i, j, k, \ell\}| = 3$, then either $k \in \{i, j\}$ or $\ell \in \{i, j\}$. These two cases contribute the same by antisymmetry, so we calculate for $k \in \{i, j\}$. In that case either $k = i$, in which case $\mathbb{E}[\mathbf{1}_{\sigma(i) > \sigma(j)} \mathbf{1}_{\sigma(i) > \sigma(\ell)}] = 1/3$, or $k = j$, in which case $\mathbb{E}[\mathbf{1}_{\sigma(i) > \sigma(j)} \mathbf{1}_{\sigma(j) > \sigma(\ell)}] = 1/6$. This yields the second and third terms.
- If $|\{i, j, k, l\}| = 2$, then we have either $(i, j) = (k, l)$, for which $\mathbb{E}[\mathbf{1}_{\sigma(i) > \sigma(j)} \mathbf{1}_{\sigma(i) > \sigma(j)}] = 1/2$, or $(i, j) = (l, k)$, for which $\mathbb{E}[\mathbf{1}_{\sigma(i) > \sigma(j)} \mathbf{1}_{\sigma(j) > \sigma(i)}] = 0$. This yields the last term.

Because A is antisymmetric, the sum over $|\{i, j, k, l\}| = 4$ is zero. Dropping this term, using antisymmetry to combine the two terms with $|\{i, j, \ell\}| = 3$, and observing that $\sum_{|\{i,j\}|=2} A_{i,j}A_{i,j} = \|A\|_F^2$, we have

$$\mathbb{E}_\sigma(\sigma \cdot A)^2 = \frac{1}{3}\left(\sum_i \left(\sum_{j \neq i} A_{i,j}\right)^2 - \|A\|_F^2\right) + \frac{1}{2}\|A\|_F^2 \geqslant \frac{1}{6}\|A\|_F^2 \tag{2.8}$$

for any antisymmetric matrix A. Combining (2.8) and (2.7) completes the proof. □

Proof of Theorem 2.1. By Lemma 2.4 it suffices to exhibit a strategy for the maximizer in the seminorm unbalancing game of Definition 2.7 that enforces $\|M - M^T\|_F \geqslant d$ after d rounds. This is rather easy. We may accomplish it by focusing only on two entries of M: the maximizer only tries to control the $1, r$ and $2, r$ entries. If in the kth round the minimizer chooses v with $v_r > 0$, the maximizer sets $a_k = 1$; else, the maximizer sets $a_k = 2$. Crucially, the entries of v are odd numbers; in particular, they are greater than 1 in absolute value. Further, all but the first and second rows of M are zero throughout the game. Thus, in the dth round, $|(M - M^T)_{2,r}| + |(M - M^T)_{1,r}| \geqslant d$ and, by antisymmetry, $|(M - M^T)_{r,2}| + |(M - M^T)_{r,1}| \geqslant d$ so $\|M - M^T\|_F \geqslant d$. □

Remark 2.2 (improving the lower bound for higher values of *r*). To prove Conjecture 1.3, it suffices to show that the maximizer can achieve $\|M\|_{S_r} = f(r)d$, where $f(r) = \omega(\log r)$. A promising strategy is to replace $\|\cdot\|_{S_r}$ with another seminorm $\|\cdot\|_*$ and show that the maximizer can enforce $\|\cdot\|_* \geqslant f(r)\|\mathrm{Id}\|_{S_r \to *}$, where Id is the identity map on $\mathrm{Mat}_{r\times r}(\mathbb{R})$. Obvious candidates such as $\|M - M^T\|_F$ and $\|M - M^T\|_1$ do not suffice. Here $\|B\|_1$ is the sum of the absolute values of entries of B. For instance, the minimizer can enforce $\|M - M^T\|_F = O(d\sqrt{\log r})$ or $\|M - M^T\|_1 = O(\sqrt{r}d)$, and even antisymmetric matrices A can achieve

$$\|A\|_{S_r} \leqslant \|A\|_F \quad \text{and} \quad \|A\|_{S_r} \leqslant \frac{\sqrt{\log r}}{\sqrt{r}}\|A\|_1.$$

The first inequality is very easy to achieve, and a result of Erdős and Moon [1] shows that the second is achieved by random ± 1 antisymmetric matrices. By the inapproximability result mentioned in Remark 2.1, it is not likely that any of the easy-to-compute norms $\|\cdot\|_*$ have both $\|\mathrm{Id}\|_{S_r \to *}$ and $\|\mathrm{Id}\|_{* \to S_r}$ bounded by constants independent of r. A candidate seminorm is the cut-norm of the top-right $(r/3) \times (2r/3)$ submatrix of the $r \times r$ matrix M: it is not hard to see that this seminorm is a lower bound for $\|M\|_{S_r}$.

3. Root-mean-squared discrepancy of permutation families

This section is concerned with upper and lower bounds for hereditary root-mean-squared discrepancy. The lower bound is Theorem 1.4, proved in Section 3.1. Afterwards, in Section 3.2, we discuss two upper bounds for which Theorem 1.4 is a tight example.

3.1 Root-mean-squared discrepancy lower bound

We now proceed with the proof of Theorem 1.4, which follows immediately from the next theorem.

Theorem 3.1 (root-mean-squared discrepancy of a 6-permutation family).

$$\mathrm{disc}_2\left([3]^d, \mathcal{A}_{S_3}\right) = \Omega(\sqrt{d}).$$

https://doi.org/10.1017/S0963548320000474 Published online by Cambridge University Press

We now give a proof strategy which motivates the rephrasing of Theorem 3.1 as a lemma about a martingale (Lemma 3.1). Fix a colouring $\chi\colon [3]^d \to 2\mathbb{Z}-1$. We must show $\operatorname{disc}_2(\mathcal{A}_{S_3},\chi)^2 = \Omega(d)$. By Lemma 2.4 and (2.3),

$$\operatorname{disc}_2(\mathcal{A}_{S_3},\chi)^2 = \mathbb{E}_{\boldsymbol{a},\sigma}[|\chi(<_\sigma \boldsymbol{a})|^2] \geqslant \frac{1}{2\sqrt{6}}\mathbb{E}_{\boldsymbol{a}}\|M_\chi(\boldsymbol{a}) - M_\chi(\boldsymbol{a})^T\|_F^2. \tag{3.1}$$

Consider again the seminorm unbalancing game of Definition 2.7. We construct a martingale by allowing the maximizer to choose rows randomly.

Definition 3.1 (martingale Y_i). Let $(M_i\colon i\in[d])$ be a joint matrix-valued random variable determined by the minimizer playing strategy χ against the maximizer choosing the sequence of rows $\boldsymbol{a}$ uniformly at random, or equivalently $M_i = M_\chi(\boldsymbol{a}[i])$. Thus

$$\mathbb{E}_{\boldsymbol{a}}\|M_\chi(\boldsymbol{a}) - M_\chi(\boldsymbol{a})^T\|_F^2 = \mathbb{E}_{\boldsymbol{a}}\|M_d - M_d^T\|_F^2.$$

It is enough to show $\mathbb{E}_{\boldsymbol{a}}\|M_d - M_d^T\|_F^2 = \Omega(d)$. Consider the sequence of random variables

$$Y_i = (M_i - M_i^T)_{1,2} + (M_i - M_i^T)_{2,3} - (M_i - M_i^T)_{1,3}.$$

By the Cauchy–Schwarz inequality, $\|M_d - M_d^T\|_F^2 \geqslant |Y_d|^2/3$, so it is enough to show that $\mathbb{E}_{\boldsymbol{a}}[Y_d^2] = \Omega(d)$. We will instead show the following, which implies Theorem 3.1 by the same reasoning.

Lemma 3.1. *Let $\chi\colon[3]^d \to 2\mathbb{Z}-1$. If $\operatorname{disc}_2(\mathcal{A}_{S_3},\chi) \leqslant 0.2(1.9/\sqrt{3})^d$ then $\mathbb{E}_{\boldsymbol{a}}[Y_d^2] \geqslant 10^{-5}d$.*

We now make a few observations from which Lemma 3.1 will follow immediately. The sequence Y_i is indeed a martingale with respect to M_i, because $Y_i - Y_{i-1} \mid M_{i-1}$ is equally likely to be $v_2 - v_3$, $v_3 - v_1$, or $v_1 - v_2$ if the minimizer chooses v in round i. Because Y_i is a martingale,

$$\mathbb{E}_{\boldsymbol{a}} Y_d^2 = \sum_{i=1}^{d} \mathbb{E}_{\boldsymbol{a}[i-1]}\left[\frac{(v_2-v_3)^2 + (v_1-v_3)^2 + (v_1-v_2)^2}{3} \,\middle|\, \boldsymbol{a}[i-1]\right]. \tag{3.2}$$

There are strategies for the minimizer that make the above quantity small, but we claim they are bad strategies if they come from a colouring χ. If $(v_2-v_3)^2 + (v_1-v_3)^2 + (v_1-v_2)^2$ is small, then v_1, v_2, v_3 are typically equal. However, strategies induced by χ satisfy that $v_1^k + v_2^k + v_3^k = v_{a_{k-1}}^{k-1}$ if the minimizer chose v^{k-1}, v^k in round $k-1, k$, respectively, and the maximizer chose a_{k-1} in round $k-1$. Further, v^k typically having equal entries should lead to the entries of v^k exponentially decreasing with k, which means that for k small they must be very large. This leads to a high discrepancy.

We now make this intuition precise.

Observation 3.2. *Let $\chi\colon[3]^d \to 2\mathbb{Z}-1$. If $\operatorname{disc}_2(\mathcal{A}_{S_3},\chi)^2 \leqslant 1.9^{2d}/(24\cdot 3^d)$, then $|\chi(\boldsymbol{\varepsilon})| \leqslant 1.9^d$.*

Proof. First, because $|\mathcal{A}_{S_3}| \leqslant 6\cdot 3^d$, if $\operatorname{disc}_2(\mathcal{A}_{S_3},\chi)^2 \leqslant 0.25\cdot 1.9^{2d}/(6\cdot 3^d)$, then $|\chi(E)| \leqslant 0.5\cdot 1.9^d$ for *every* $E\in\mathcal{A}_{S_3}$. To finish, we need only express $|\chi(\varepsilon)|$ in terms of the discrepancy of edges. Let id denote the identity permutation. By definition, $\chi(\varepsilon) = \chi(\leqslant_{\mathrm{id}} \bar{3})$, but $\leqslant_{\mathrm{id}} \bar{3} \notin \mathcal{A}_{S_3}$. However, $\leqslant_{\mathrm{id}} \bar{3}$ is the disjoint union of $<_{\mathrm{id}} \bar{3}$ and $\bar{3}$. In turn, because the vertices are ordered oppositely under id and $\tau = (1,3)$, we have $\bar{3} = <_\tau \bar{3}2$. It follows that $|\chi(\boldsymbol{\varepsilon})| = |\chi(<_{\mathrm{id}} \bar{3}) + \chi(<_\tau \bar{3}2)| \leqslant 1.9^d$. □

Next we show that the assumption $|\chi(\boldsymbol{\varepsilon})| \leqslant 1.9^d$ implies many cancellations, and that this implies (3.2) is large. For $\boldsymbol{a}\in[r]^0\cup[r]^1\cup\cdots\cup[r]^d$, define the *cancellation* of χ at $\boldsymbol{a}$ by

$$C_\chi(\boldsymbol{a}) = \sum_{i\in[3]} |\chi(\boldsymbol{a}i)| - |\chi(\boldsymbol{a})|. \tag{3.3}$$

https://doi.org/10.1017/S0963548320000474 Published online by Cambridge University Press

For $i \in \{0, \dots, d-1\}$, define the average cancellation $\overline{C^i_\chi} = \mathbb{E}_{\boldsymbol{a}\in[r]^i} C_\chi(\boldsymbol{a})$. The following two propositions, along with Observation 3.2, imply Lemma 3.1.

Proposition 3.2. *Let* $\chi\colon [3]^d \to 2\mathbb{Z}-1$. *Then*

$$\mathbb{E}Y_d^2 \geqslant \frac{1}{9d}\left(\sum_{i=0}^{d-1} \overline{C^i_\chi}\right)^2.$$

Proposition 3.3. *Let* $\chi\colon [3]^d \to 2\mathbb{Z}-1$. *If* $|\chi(\boldsymbol{\varepsilon})| \leqslant 1.9^d$ *and* $d \geqslant 400$, *then*

$$\sum_{i=0}^{d-1} \overline{C^i_\chi} \geqslant 0.01d.$$

Proof of Proposition 3.2. In response to $\boldsymbol{a} = a_1 \dots a_{k-1}$, the maximizer plays the vector $v = (\chi(\boldsymbol{a}1), \chi(\boldsymbol{a}2), \chi(\boldsymbol{a}3))$. Then

$$C_\chi(\boldsymbol{a})^2 = (|v_1| + |v_2| + |v_3| - |v_1 + v_2 + v_3|)^2 \tag{3.4}$$
$$\leqslant (|v_1 - v_2| + |v_2 - v_3| + |v_3 - v_1|)^2 \tag{3.5}$$
$$\leqslant 3(|v_1 - v_2|^2 + |v_2 - v_3|^2 + |v_3 - v_1|^2). \tag{3.6}$$

Here (3.5) is the inequality $|a| + |b| + |c| - |a+b+c| \leqslant |a-b| + |b-c| + |c-a|$, which can be proved by cases: without loss of generality, $a \leqslant b \leqslant c$, if all are positive, then both sides vanish; else, without loss of generality $a \leqslant 0 \leqslant b$. In this case the left-hand side is $2|a|$ but the right-hand side is $2|a| + 2|c|$. Thus, if the strategy of the minimizer is induced by χ, using (3.6) and (3.2) we have

$$\begin{aligned}
\mathbb{E}_{\boldsymbol{a}} Y_d^2 &= \sum_{i=1}^{d} \mathbb{E}_{\boldsymbol{a}[i-1]}\left[\frac{(v_2-v_3)^2 + (v_1-v_3)^2 + (v_1-v_2)^2}{3} \,\middle|\, \boldsymbol{a}[i-1]\right] \\
&\geqslant \frac{1}{9}\sum_{i=1}^{d} \mathbb{E}_{\boldsymbol{a}\in[r]^{i-1}}[C_\chi(\boldsymbol{a})^2] \\
&\geqslant \frac{1}{9}\sum_{i=0}^{d-1} \overline{C^i_\chi}^2 \\
&\geqslant \frac{1}{9d}\left(\sum_{i=0}^{d-1} \overline{C^i_\chi}\right)^2.
\end{aligned}$$

□

Proof of Proposition 3.3. Define the average absolute value

$$\overline{|\chi_i|} = \mathbb{E}_{\boldsymbol{a}\in[r]^i} |\chi(\boldsymbol{a})|.$$

Note that $\overline{|\chi_i|} \geqslant 1$. Thus there exists $j \in \{1, \dots, \lceil 0.99d \rceil\}$ such that $\overline{|\chi_{j-1}|} \leqslant 2\overline{|\chi_j|}$, else

$$|\chi(\boldsymbol{\varepsilon})| = \overline{|\chi_0|} \geqslant 2^{0.99d} > 1.9^d.$$

Taking the expectation of both sides of the definition (3.3) of cancellation yields the identity

$$\overline{C^i_\chi} = 3\overline{|\chi_{i+1}|} - \overline{|\chi_i|},$$

so

$$\sum_{i=j-1}^{d-1} \overline{C_\chi^i} = 3\overline{|\chi_d|} - \overline{|\chi_{j-1}|} + 2\sum_{i=j}^{d-1} \overline{|\chi_i|} \geqslant 2\sum_{i=j+1}^{d-1} \overline{|\chi_i|} \geqslant 2(\lfloor 0.01d \rfloor - 2).$$

The right-hand side is at least $0.01d$ provided d is at least 400. □

3.2 Root-mean-squared discrepancy upper bounds

In this section we prove two upper bounds, Theorem 3.4 and Theorem 3.5, for the hereditary root-mean-squared discrepancy. The bounds are analogous to the following bound for hereditary ℓ_∞-discrepancy from [5].

Theorem 3.3 **([5]).** *Let $m = |\mathcal{A}|$. Then*

$$\text{herdisc}_\infty(\Omega, \mathcal{A}_1 + \cdots + \mathcal{A}_k) = O\left(\sqrt{k}\log(mn)\sqrt{\log n}\max_{i\in[k]} \text{herdisc}_\infty(\Omega, \mathcal{A}_i)\right),$$

where + denotes the multiset sum (union with multiplicity).

The first is a bound for the hereditary root-mean-squared discrepancy of a union of families with bounded hereditary discrepancy, and follows relatively straightforwardly from the results of [5], [8] and [3].

Theorem 3.4 **(communicated by Aleksandar Nikolov).** *We have*

$$\text{herdisc}_2(\Omega, \mathcal{A}_1 + \cdots + \mathcal{A}_k) = O\left(\sqrt{k\log n}\max_{i\in[k]} \text{herdisc}_\infty(\Omega, \mathcal{A}_i)\right).$$

The next result bounds the same quantity, but for unions of families with bounded *root-mean-squared* discrepancy. The proof mainly relies on a result appearing in [3] and [8], but requires one technical lemma.

Theorem 3.5. *We have*

$$\text{herdisc}_2(\Omega, \mathcal{A}_1 + \cdots + \mathcal{A}_k) = O\left(k\sqrt{\log n}\max_{i\in[k]} \text{herdisc}_2(\Omega, \mathcal{A}_i)\right).$$

If $(\Omega, \mathcal{A})$ is a 1-permutation family, then $\text{herdisc}_\infty(\Omega, \mathcal{A}) = 1$. Combined with Theorem 3.4, we immediately recover the bound which was proved directly in [9].

Corollary 3.4 **([9]).** *If $(\Omega, \mathcal{A})$ is a k-permutation family, then* $\text{herdisc}_2(\Omega, \mathcal{A}) \leqslant \sqrt{k\log n}$.

Theorem 1.4 implies that, for constant k, Theorem 3.4, Theorem 3.5 and hence Corollary 3.4 are tight. It would be interesting to improve k to $\sqrt{k}$ in Theorem 3.5, thereby providing a common strengthening of Theorem 3.4 and Theorem 3.5.

To prove Theorem 3.4 and Theorem 3.5 we introduce three quantities. The first is an approximation for hereditary discrepancy, and the latter two are approximations for hereditary root-mean-squared discrepancy.

Definition 3.2 **(lower bounds).** Let A denote the $|\Omega| \times |\mathcal{A}|$ incidence matrix of $(\Omega, \mathcal{A})$.

https://doi.org/10.1017/S0963548320000474 Published online by Cambridge University Press

(1) Define

$$\operatorname{detlb}(\Omega, \mathcal{A}) = \max_k \max_B |\det(B)|^{1/k},$$

where B runs over all $k \times k$ submatrices of A.

(2) Define

$$\operatorname{detlb}_2(\Omega, \mathcal{A}) = \max_{\Gamma \subset \Omega} \sqrt{\frac{m|\Gamma|}{8\pi e}} \det(A|_\Gamma^T A|_\Gamma)^{1/(2|\Gamma|)}.$$

(3) Let λ_l be the lth largest eigenvalue of $A^T A$. Define

$$\operatorname{kgl}(\Omega, \mathcal{A}) = \max_{1 \leqslant l \leqslant \min\{|\Omega|, |\mathcal{A}|\}} \frac{l}{e} \sqrt{\frac{\lambda_l}{8\pi |\Omega||\mathcal{A}|}} \quad \text{and} \quad \operatorname{herkgl}(\Omega, \mathcal{A}) = \max_{\Gamma \subset \Omega} \operatorname{kgl}(\Gamma, \mathcal{A}|_\Gamma).$$

We now state the bounds obtained from each of the three quantities, beginning with the first. The lower bound in the next theorem is from [4], and the upper bound (which we will not need here) is from [5].

Theorem 3.6 **([4, 5]).** *Let $n = |\Omega|$ and $m = |\mathcal{A}|$. Then*

$$\frac{1}{2} \operatorname{detlb}(\Omega, \mathcal{A}) \leqslant \operatorname{herdisc}_\infty(\Omega, \mathcal{A}) = O(\log(mn)\sqrt{\log n} \cdot \operatorname{detlb}(\Omega, \mathcal{A}))$$

The next theorem shows how the second two quantities in Definition 3.2 approximate discrepancy. The upper bound follows from Corollary 2 and the proof of Theorem 7 in [3]. The lower bounds are from Theorem 6 in [3], but up to a constant the middle inequality is a corollary of Theorem 11 in the earlier work [8].

Theorem 3.7 **([3, 8]).** *We have*

$$\operatorname{herkgl}(\Omega, \mathcal{A}) \leqslant \operatorname{detlb}_2(\Omega, \mathcal{A}) \leqslant \operatorname{herdisc}_2(\Omega, \mathcal{A}) = O(\sqrt{\log n} \cdot \operatorname{herkgl}(\Omega, \mathcal{A})).$$

Finally we will need to relate detlb and detlb_2. Applying the Cauchy–Binet identity to $\det(A^T A)$ implies

$$\operatorname{detlb}_2(\Omega, \mathcal{A}) = O(\operatorname{detlb}(\Omega, \mathcal{A})). \tag{3.7}$$

Both of the quantities detlb and herkgl behave well under unions; the former was shown in [5], and the latter we show here.

Theorem 3.8 **([5]).** *We have*

$$\operatorname{detlb}(\Omega, \mathcal{A}_1 + \cdots + \mathcal{A}_k) = O\left(\sqrt{k} \max_{i \in [k]} \operatorname{detlb}(\Omega, \mathcal{A}_i)\right).$$

Theorem 3.9. *We have*

$$\operatorname{herkgl}(\Omega, \mathcal{A}_1 + \cdots + \mathcal{A}_k) \leqslant k \max_{i \in [k]} \operatorname{herkgl}(\Omega, \mathcal{A}_i).$$

Proof. Let $C = \max_{i \in [k]} \operatorname{herkgl}(\Omega, \mathcal{A}_i)$. It is enough to show $\operatorname{kgl}(\Gamma, (\mathcal{A}_1 + \cdots + \mathcal{A}_k)|_\Gamma) \leqslant kC$ for any $\Gamma \subset \Omega$. Let $|\Gamma| = n$, $m_i = |\mathcal{A}_i|$, and $\sum m_i = m$. If A_i is the incidence matrix of $(\Gamma, \mathcal{A}_i|_\Gamma)$ and A that of $(\Gamma, (\mathcal{A}_1 + \cdots + \mathcal{A}_k)|_\Gamma)$, then

$$A^T A = A_i^T A_i + \cdots + A_i^T A_i.$$

Weyl's inequality on the eigenvalues of Hermitian matrices asserts that if H_1 and H_2 are $n \times n$ Hermitian matrices then $\lambda_{i+j-1}(H_1 + H_2) \leqslant \lambda(H_1)_i + \lambda(H_2)_j$ for all $1 \leqslant i, j \leqslant i + j - 1 \leqslant n$. Applying this inequality inductively,

$$\lambda_l(A^T A) \leqslant \sum_{i=1}^{k} \lambda_{\lceil l/k \rceil}(A_i^T A_i).$$

Thus

$$\begin{aligned} \operatorname{kgl}(\Gamma, (\mathcal{A}_1 + \cdots + \mathcal{A}_k)|_\Gamma) &= \max_{1 \leqslant l \leqslant \min\{n,m\}} \frac{l}{e}\sqrt{\frac{\lambda_l(A^T A)}{8\pi mn}} \\ &\leqslant \max_{1 \leqslant l \leqslant \min\{n,mk\}} \frac{l}{e}\sqrt{\frac{\sum_{i=1}^{k} \lambda_{\lceil l/k \rceil}(A_i^T A_i)}{8\pi mn}} \\ &\leqslant kC, \end{aligned}$$

where in the last line we used $\sum m_i = m$ and $\lambda_{\lceil l/k \rceil}(A_i^T A_i) \leqslant 8\pi m_i n (Cek/l)^2$ from our assumption that $\operatorname{kgl}(\Gamma, \mathcal{A}_i|_\Gamma) \leqslant \operatorname{herkgl}(\Omega, \mathcal{A}_i) \leqslant C$. □

Theorem 3.5 is immediate from Theorem 3.9 and Theorem 3.7. We now prove Theorem 3.4.

Proof of Theorem 3.4. By Theorem 3.7, (3.7), Theorem 3.8 and Theorem 3.6,

$$\begin{aligned} \operatorname{herdisc}_2(\Omega, \mathcal{A}_1 + \cdots + \mathcal{A}_k) &= O(\sqrt{\log n}\, \operatorname{detlb}_2(\Omega, \mathcal{A}_1 + \cdots + \mathcal{A}_k)) \\ &= O(\sqrt{\log n}\, \operatorname{detlb}(\Omega, \mathcal{A}_1 + \cdots + \mathcal{A}_k)) \\ &= O\Big(\sqrt{k \log n} \max_{i \in [k]} \operatorname{detlb}(\Omega, \mathcal{A}_i)\Big) \\ &= O\Big(\sqrt{k \log n} \max_{i \in [k]} \operatorname{herdisc}_\infty(\Omega, \mathcal{A}_i)\Big). \end{aligned}$$ □

Theorem 1.4 also shows tightness for Theorem 3.7. By (3.7) and Theorem 3.8 we have that $\operatorname{detlb}_2(\Omega, \mathcal{A})$ is constant for k-permutation families $(\Omega, \mathcal{A})$ with constant k. Thus Theorem 1.4 shows that Theorem 3.7 is best possible in the sense that there can be a $\Omega(\sqrt{\log n})$ gap between $\operatorname{detlb}_2(\Omega, \mathcal{A})$ and $\operatorname{herdisc}_2(\Omega, \mathcal{A})$.

Acknowledgements

The author would like to thank Michael Saks and Aleksandar Nikolov for many interesting discussions. The author also thanks Aleksandar Nikolov and Shachar Lovett for suggesting the application of this argument to the root-mean-squared discrepancy, and especially to Aleksandar Nikolov for communicating Theorem 3.4 and suggesting the connection with [3], [8] and [5].

References

[1] Erdős, P. and Moon, J. (1965) On sets of consistent arcs in a tournament. *Canad. Math. Bull.* **8** 269–271.

[2] Guruswami, V., Manokaran, R. and Raghavendra, P. (2008) Beating the random ordering is hard: inapproximability of maximum acyclic subgraph. In *IEEE 49th Annual Symposium on Foundations of Computer Science (FOCS '08)*, pp. 573–582. IEEE.

[3] Larsen, K. G. (2019) Constructive discrepancy minimization with hereditary L2 guarantees. In *36th International Symposium on Theoretical Aspects of Computer Science (STACS 2019)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.

[4] Lovász, L., Spencer, J. and Vesztergombi, K. (1986) Discrepancy of set-systems and matrices. *European J. Combin.* **7** 151–160.
[5] Matoušek, J. (2013) The determinant bound for discrepancy is almost tight. *Proc. Amer. Math. Soc.* **141** 451–460.
[6] Newman, A., Neiman, O. and Nikolov, A. (2012) Beck's three permutations conjecture: a counterexample and some consequences. In *IEEE 53rd Annual Symposium on Foundations of Computer Science (FOCS)*, pp. 253–262. IEEE.
[7] Newman, A. and Nikolov, A. (2011) A counterexample to Beck's conjecture on the discrepancy of three permutations. arXiv:1104.2922
[8] Nikolov, A., Talwar, K. and Zhang, L. (2013) The geometry of differential privacy: the sparse and approximate cases. In *45th Annual ACM Symposium on Theory of Computing (STOC '13)*, pp. 351–360. ACM.
[9] Spencer, J. H. (1987) *Ten Lectures on the Probabilistic Method*, Vol. 52 of Proc. CBMS-NRM Regional Conference Series in Applied Mathematics. SIAM.
[10] Spencer, J. H., Srinivasan, A. and Tetali, P. (2018) The discrepancy of permutation families.

Cite this article: Franks C (2021). A simplified disproof of Beck's three permutations conjecture and an application to root-mean-squared discrepancy. *Combinatorics, Probability and Computing* **30**, 398–411. https://doi.org/10.1017/S0963548320000474