

Level sets and maximum likelihood estimation for the Ising model

Tomasz Skalski ^{*1} and Tomasz Stroiński ^{†1}

¹Faculty of Pure and Applied Mathematics, Wrocław University of Science and Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland

November 27, 2025

Abstract

Bogdan et al. established a new criterion to determine the existence of a maximum likelihood estimator in discrete exponential families. It uses the notion of the set of uniqueness, which allows to apply the problem to the Ising model from statistical mechanics. We propose a full characterization of the existence of the MLE in the Ising model among the level sets used in related combinatorial problems. Then we establish new bounds for the size of the smallest set of uniqueness for the products of Rademacher functions.

Key words: maximum likelihood, discrete exponential family, Ising model.

Mathematics Subject Classification (2010): 62H12.

1 Introduction

The class of exponential families is one of the most popular classes of distributions in probability and statistics. Apart from many properties that make them useful to derive better theoretical results and apply in the practical setting [17, Sec. 2.7], [16], they allow an easier analysis of the Maximum Likelihood Estimator (MLE). However, in some cases the MLE may not exist, as pointed out, e.g., by Crain [8, 9] in situations, where the number of parameters is too large. Barndorff-Nielsen [3, Theorem 9.13] established a full characterization of the MLE existence for exponential families, however, in general it is rather difficult to apply in practice [5]. Bogdan et al. [5] proposed a new, straightforward criterion of the MLE existence for discrete exponential families, that was inspired by the article by Bogdan and Bogdan [4] in the continuous case. This criterion uses the notion of the set of uniqueness, i.e., the set that enforces that the function from a given class that vanishes on such set is the constant zero function on the whole space. The article by Bogdan et al. [5] enriches the discussion on the existence of the MLE with more exact results on the Rademacher functions and their products, both of which are related to the discrete hypercube $\mathcal{X} = \{-1, +1\}^k$. These results include the theoretical conditions for the MLE existence for the given sample, as well as the probability bounds for the sample size needed to ensure it. However, in general, for the products of Rademacher functions, establishing the exact conditions to ensure the existence of the MLE remains the open problem.

In this article we focus on the fact that in many cases of the products of Rademacher functions, the minimal samples, for which the MLE exists, can be described as the level sets, i.e., the sets of those points in $\mathcal{X} = \{-1, +1\}^k$ having the number of positive coordinates that belongs to a given set of numbers. The level sets have been used, e.g., for establishing the smallest subsets of \mathcal{X} that intersect with every subcube of the given size [1], [2, Sec. 2.5], [12]. This problem is also known as the subcube allocation problem [12]. It can be used, e.g., in the analysis of the hypercube processors, known from their interconnection properties, their fault-tolerance as well as the ability to support multiple users, that are assigned the subcubes of given size, cf. [11].

^{*}corresponding author, email: Tomasz.Skalski@pwr.edu.pl

[†]email: Tomasz.Stroinski@pwr.edu.pl

The paper is composed as follows. In Section 2 we include the full characterization of the existence of the MLE for the level sets in the Ising model. Section 3 we discuss the bounds on the size of the smallest set of uniqueness for the products of Rademacher functions, improving the upper bounds. Section 4 covers the proofs from Section 2.

1.1 Discrete exponential families

For $k \in \mathbb{N}$, let us consider $\mathcal{X} := \{-1, +1\}^k$, the k -dimensional discrete cube with the uniform weight $\mu(x) = 2^{-k}$, $x \in \mathcal{X}$. Thus, $K = |\mathcal{X}| = 2^k$. We adopt the convention of [5] that $\mathbb{R}^{\mathcal{X}}$ is the class of all the real functions on \mathcal{X} . In further discussions, as \mathcal{B} we denote a fixed linear subspace of $\mathbb{R}^{\mathcal{X}}$ containing the constant function $\varphi \equiv 1$.

We often refer to the cone of non-negative functions in \mathcal{B} :

$$\mathcal{B}_+ := \{\phi \in \mathcal{B} : \phi \geq 0\}.$$

We define, as in [5], the exponential family

$$e(\mathcal{B}) = \{e(r) : r \in \mathcal{B}\}.$$

As \mathcal{X} is a finite set, $e(\mathcal{B})$ is known as a discrete exponential family, cf. [5]. Let $x_1, \dots, x_n \in \mathcal{X}$ be an iid sample. For an exponential density p , its likelihood function $L_p(x_1, \dots, x_n)$ is defined as $\prod_{i=1}^n p(x_i)$ and its log-likelihood function $l_p(x_1, \dots, x_n)$ is defined as $\log L_p(x_1, \dots, x_n)$.

An element \hat{p} of $e(\mathcal{B})$ is called a maximum likelihood estimator (MLE) for x_1, \dots, x_n and $e(\mathcal{B})$ if

$$L_{\hat{p}}(x_1, \dots, x_n) = \sup_{p \in e(\mathcal{B})} L_p(x_1, \dots, x_n),$$

or, equivalently,

$$l_{\hat{p}}(x_1, \dots, x_n) = \sup_{p \in e(\mathcal{B})} l_p(x_1, \dots, x_n).$$

To establish new elegant results on the existence of the MLE for the exponential family $e(\mathcal{B})$, we use the notion of the set of uniqueness from [5].

Definition 1.1. [5, Sec. 2.] Let $U \subset \mathcal{X}$. Let \mathcal{B} be a fixed linear subspace of $\mathbb{R}^{\mathcal{X}}$. We say that U is a set of uniqueness for \mathcal{B} if $\phi = 0$ is the only function in \mathcal{B} such that $\phi = 0$ on U . Similarly, we say that U is a set of uniqueness for \mathcal{B}_+ if $\phi = 0$ is the only function in \mathcal{B}_+ such that $\phi = 0$ on U .

Remark 1.2. [5, Sec. 2.] If U is a set of uniqueness for \mathcal{B} , then U is a set of uniqueness for \mathcal{B}_+ , since $\mathcal{B}_+ \subset \mathcal{B}$.

1.2 Rademacher functions

Motivated by the characterization of the existence of the MLE from [5], we analyze the Rademacher functions as well as the products of two Rademacher functions, better known as the Ising model.

For $j = 1, \dots, k$ and $x = (x_1, \dots, x_k) \in \mathcal{X}$ we define Rademacher functions:

$$r_j(x) = x_j,$$

and we denote $r_0(x) = 1$. Let

$$\mathcal{B}^k = \text{Lin}\{r_0, r_1, \dots, r_k\}.$$

We define the $(k - q)$ -subcube of \mathcal{X} by fixing the values of exactly q coordinates:

$$(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_k) \in \{-1, +1, *\}^k,$$

where $*$ denotes for both values of a given coordinate, i.e., $* = \{-1, +1\}$. For example, for $k = 4$, the subcube defined by fixing the first coordinate at $+1$ and the last coordinate at -1 is denoted as $(+1, *, *, -1)$.

With the notion of the $(k - q)$ -subcube we discuss the spaces spanned by products of q Rademacher functions. Let $k \in \mathbb{N}$ and $1 \leq q \leq k$. We define

$$\mathcal{B}_q^k = \text{Lin}\{w_L : L \subset \{1, \dots, k\} \text{ and } |L| \leq q\},$$

where

$$w_L(x) = \prod_{j \in L} r_j(x), \quad x \in \{-1, +1\}^k, \quad L \subset \{1, \dots, k\},$$

are the Walsh functions, see, e.g., Bogdan et al. [5] or Jendrej et al. [13].

Similarly, we define the non-negative cone of functions in \mathcal{B}_q^k as $(\mathcal{B}_q^k)_+$.

Remark 1.3. [5, Remark 5.6] *Let $1 \leq q_1 \leq q_2 \leq k$. Then every set U of uniqueness for $(\mathcal{B}_{q_2}^k)_+$ is of uniqueness for $(\mathcal{B}_{q_1}^k)_+$, since $(\mathcal{B}_{q_1}^k)_+ \subset (\mathcal{B}_{q_2}^k)_+$. Analogously, every set U of uniqueness for $\mathcal{B}_{q_2}^k$ is of uniqueness for $\mathcal{B}_{q_1}^k$, since $\mathcal{B}_{q_1}^k \subset \mathcal{B}_{q_2}^k$.*

In this article, we primarily focus on the space spanned by the products of two Rademacher functions, \mathcal{B}_2^k , and its non-negative cone, $(\mathcal{B}_2^k)_+$, which is in fact the Ising model described below.

1.3 Ising model

Let $G = (V, E)$ be an undirected graph, where $V = \{1, 2, \dots, k\}$ is a set of its vertices and $E \subset \binom{V}{2} = \{(i, j) : 1 \leq i < j \leq k\}$ is a set of its edges.

The Ising model was proposed by Lenz in 1920 and developed by his student Ernst Ising in 1925 [7]. It is defined with Boltzmann distribution by the following Hamiltonian [10]:

Definition 1.4. *Let $x = \{x_i : i \in V\}$, where $x_i \in \{-1, +1\}$. Then the Boltzmann measure is equal to:*

$$\mu(x) = \frac{1}{Z(B, \beta)} \exp \left(B \sum_{i \in V} x_i + \beta \sum_{(i, j) \in E} x_i x_j \right),$$

where $Z(B, \beta)$ is the normalizing constant (partition function). The parameters B and β can be interpreted respectively as the magnetic field and the inverse temperature.

In our case, we assume the graph $G = (V, E)$ to be a complete graph consisting of k vertices. Thus, each x_i is a point in $\{-1, +1\}^k$, every vertex refers to a single Rademacher function and each pair of vertices (i.e. each edge) refers to a product of two different Rademacher functions. The constant function $f \equiv 1$ is implicitly hidden by the partition function.

The discrete exponential family approach to the Ising model was applied, e.g., by Johnson et al. [14, Sec. 2.2], Wainwright and Jordan [18, Example 3.1.] and Wang et al. [19]. Wang et al. [19] treat the Ising model as a subclass of discrete hierarchical models obtained by choosing a set of rows from the Walsh matrix, where each row corresponds to the product of Rademacher functions r_i for every i from $I \subset \{1, \dots, k\}$. For the Ising model, they choose the rows related to the products of at most two Rademacher functions, which generate exactly the $(\mathcal{B}_2^k)_+$ space, cf. [5, Sec. 5].

2 Level sets

2.1 Introduction and motivation

Since $|\mathcal{X}| = |\{-1, +1\}^k| = 2^k$, there are exactly 2^{2^k} subsets of \mathcal{X} . Thus, establishing all sets of uniqueness having the smallest cardinality becomes a challenging problem.

Therefore, we consider a specific class of subsets of $\mathcal{X} = \{-1, +1\}^k$, called level sets. For a given $x \in \mathcal{X}$ and $D \subset \{0, 1, \dots, k\}$, the level set $W_D^k(x)$ is defined as the set of all points that differ from x at exactly d coordinates, where $d \in D$. To be more specific, for x and D we define the level set as:

$$W_D^k(x) = \left\{ y \in \{-1, +1\}^k : \text{dist}(x, y) \in D \right\},$$

where dist is the Hamming distance, i.e. $\text{dist}(x, y) := \#\{i : x_i \neq y_i\}$.

Without loss of generality, we denote W_D^k by $W_D^k((-1, \dots, -1))$, which is isomorphic to $W_D^k(x)$ for every x . When k is known from the context, we simplify the notation to W_D .

The optimality of level sets appears in a merely related problem, that is a d -chromatic coloring of $\mathcal{X} = \{-1, +1\}^k$ in such way that every d -subcube claims every color in \mathcal{X} , where d is defined as $k - q$. As proven in [1, Theorem 7], the asymptotically maximal number of colors, for which such coloring is admissible, equals $d + 1$. The proof is algorithmic and shows that the optimal distribution of $(d + 1)$ colors onto the vertices of \mathcal{X} is to give the color $j = 0, 1, \dots, d$ to the level set $W_{\{j, j+d+1, j+2(d+1), \dots\}}$.

In the same paper, the authors connect the d -chromatic coloring with the smallest cardinality of the subset of \mathcal{X} that intersects with its every d -subcube. In fact, they denote such cardinality as $g(k, d)$ and prove it to be not greater than $2^k/(d+1)$.

Level sets often appear as important elements of minimal sets of uniqueness for $(\mathcal{B}_q^k)_+$ with respect to inclusion, as summarized below.

Remark 2.1.

1. The only set of uniqueness for $(\mathcal{B}_k^k)_+$ is $W_{\{0,1,2,\dots,k\}} = \mathcal{X}$.
2. All minimal sets of uniqueness for $(\mathcal{B}_{k-1}^k)_+$ are isomorphic to $W_{\{0,2,4,\dots\}}$.
3. The set $W_{\{0,k\}}$ is the smallest set of uniqueness for $(\mathcal{B}_1^k)_+$.

Proof. The first statement was proven in [5, Lemma 3.1]. The second statement follows directly from [5, Theorem 5.3]. In this case, there are exactly two minimal sets of uniqueness, the first is $W_{\{0,2,\dots\}}$ and the second is $W_{\{1,3,\dots\}}$, which is isomorphic to the previous one with the transformation $x_1 \mapsto -x_1$.

Regarding the third statement, $W_{\{0,k\}}$ consists of exactly two points: $x = (-1, -1, \dots, -1)$ and $-x = (1, 1, \dots, 1)$. It is a set of uniqueness for $(\mathcal{B}_1^k)_+$, which was proven in [5, Example 3.7]. Smaller sets consist only of one or zero points, therefore none of them intersects with every half-cube of \mathcal{X} . This contradicts [5, Theorem 3.6], and therefore the sets isomorphic to $W_{\{0,k\}}$ are the only sets of uniqueness for $(\mathcal{B}_1^k)_+$ having the smallest cardinality possible. \square

To derive more complex results on the sets of uniqueness, we need a following representation.

Let \mathcal{S}_q^k be the set of all $(k-q)$ -subcubes of \mathcal{X} . That said, each function from \mathcal{B}_q^k can be denoted as

$$\varphi(x) = \sum_{S \in \mathcal{S}_q^k} \alpha_S \mathbb{1}_{\{x \in S\}}.$$

Notice that in every $(k-q)$ -subcube, the number of positive coordinates of points varies from j to $j+k-q$, where $0 \leq j \leq q$. Therefore, we perform the partition

$$\mathcal{S}_q^k = \mathcal{S}_0 \cup \mathcal{S}_1 \cup \dots \cup \mathcal{S}_q, \quad (2.1)$$

where \mathcal{S}_j is the set of subcubes in which the number of positive coordinates of points varies from j to $j+k-q$.

2.2 Full characteristics for the Ising model

We start with an useful lemma, which states that the sum of values of a function $\varphi \in \mathcal{B}_q^k$ over a level set W_D can be represented by $q+1$ coordinates T_0, T_1, \dots, T_q .

Lemma 2.2. Let $\varphi \in \mathcal{B}_q^k$ and $D \subset \{0, 1, \dots, k\}$. If $U = W_D$, then

$$\sum_{x \in U} \varphi(x) = 2^{k-q} \sum_{j=0}^q T_j, \quad (2.2)$$

where $T_j = \sum_{S \in \mathcal{S}_j} \alpha_S$ for every $j = 0, 1, \dots, q$ and $\alpha_S \in \mathbb{R}$ for every $S \in \mathcal{S}_q^k = \mathcal{S}_0 \cup \mathcal{S}_1 \cup \dots \cup \mathcal{S}_q$.

Proof.

$$\begin{aligned} \sum_{x \in \mathcal{X}} \varphi(x) &= \sum_{x \in \mathcal{X}} \left(\sum_{S \in \mathcal{S}_q^k} \alpha_S \mathbb{1}_{\{x \in S\}} \right) = \sum_{x \in \mathcal{X}} \left(\sum_{S \in \mathcal{S}_0} \alpha_S \mathbb{1}_{\{x \in S\}} \right) + \dots + \sum_{x \in \mathcal{X}} \left(\sum_{S \in \mathcal{S}_q} \alpha_S \mathbb{1}_{\{x \in S\}} \right) \\ &= \sum_{S \in \mathcal{S}_0} \left(\alpha_S \sum_{x \in \mathcal{X}} \mathbb{1}_{\{x \in S\}} \right) + \dots + \sum_{S \in \mathcal{S}_q} \left(\alpha_S \sum_{x \in \mathcal{X}} \mathbb{1}_{\{x \in S\}} \right) = \sum_{S \in \mathcal{S}_0} \alpha_S |S| + \dots + \sum_{S \in \mathcal{S}_q} \alpha_S |S| \\ &= |S| \left(\sum_{S \in \mathcal{S}_0} \alpha_S + \dots + \sum_{S \in \mathcal{S}_q} \alpha_S \right) = 2^{k-q} \left(\sum_{S \in \mathcal{S}_0} \alpha_S + \dots + \sum_{S \in \mathcal{S}_q} \alpha_S \right) = 2^{k-q} (T_0 + T_1 + \dots + T_q). \end{aligned}$$

\square

Therefore, showing that $T_0 = \dots = T_q = 0$ is sufficient to prove that $\varphi \equiv 0$.

Lemma 2.3. *Consider $(\mathcal{B}_q^k)_+$. Let $S \in S_i$. Then $|W_j \cap S| = \binom{k-q}{j-i}$, with the convention that $\binom{n}{r} = 0$ for $r < 0$ and for $r > n$.*

Proof. By definition, every point from $S \in S_i$ has q fixed coordinates, where i of them are positive and $q-i$ are negative. To obtain a point from S that has exactly j positive coordinates in total, one needs to choose exactly $j-i$ from $k-q$ remaining ones, which can be done in $\binom{k-q}{j-i}$ ways. \square

Lemma 2.4. *The set $W_{\{1,k\}}$ is a minimal set of uniqueness for $(\mathcal{B}_2^k)_+$ with respect to inclusion.*

Proof. Let $\varphi \in (\mathcal{B}_2^k)_+$. It means that φ is non-negative and can be denoted as

$$\varphi(x) = \sum_{S \in S_2^k} \alpha_S \mathbb{1}_{\{x \in S\}} = \sum_{S \in S_0} \alpha_S \mathbb{1}_{\{x \in S\}} + \sum_{S \in S_1} \alpha_S \mathbb{1}_{\{x \in S\}} + \sum_{S \in S_2} \alpha_S \mathbb{1}_{\{x \in S\}},$$

where $\alpha_S \in \mathbb{R}$ and $x \in \mathcal{X}$. As φ is non-negative, to show that $\varphi \equiv 0$ on \mathcal{X} it is sufficient to show that $\sum_{x \in \mathcal{X}} \varphi(x) = 0$. By Lemma 2.2, this is equivalent to $T_0 + T_1 + T_2 = 0$. Assume that $\varphi(U) = \{0\}$. Because $W_1 \subset U$, we obtain

$$0 = \sum_{x \in W_1} \varphi(x) = \sum_{x \in W_1} \left(\sum_{S \in S_0} \alpha_S |W_1 \cap S| + \sum_{S \in S_1} \alpha_S |W_1 \cap S| + \sum_{S \in S_2} \alpha_S |W_1 \cap S| \right) = (k-2)T_0 + T_1.$$

Moreover, $(+1, +1, \dots, +1) \in U$, thus

$$0 = \varphi((+1, +1, \dots, +1)) = \sum_{S \in S_2} \alpha_S = T_2.$$

An observation that W_0 and W_{k-1} are subsets of \mathcal{X} , implies that:

$$0 \leq \varphi((-1, -1, \dots, -1)) = \sum_{S \in S_0} \alpha_S = T_0,$$

and respectively

$$\begin{aligned} 0 &\leq \sum_{x \in W_{k-1}} \varphi(x) = \sum_{x \in W_{k-1}} \left(\sum_{S \in S_0} \alpha_S |W_{k-1} \cap S| + \sum_{S \in S_1} \alpha_S |W_{k-1} \cap S| + \sum_{S \in S_2} \alpha_S |W_{k-1} \cap S| \right) \\ &= T_1 + (k-2)T_2 = T_1. \end{aligned}$$

Thus $T_0 = T_1 = T_2 = 0$, which implies that $\varphi(X) = \{0\}$ and $U = W_{\{1,k\}}$ is a set of uniqueness for $(\mathcal{B}_2^k)_+$. \square

Our main result is the full characterization of sets of uniqueness for $(\mathcal{B}_2^k)_+$ among the level sets in $\mathcal{X} = \{-1, +1\}^k$.

Theorem 2.5. *The set W_D is of uniqueness for $(\mathcal{B}_2^k)_+$ if and only if there exist $i, j \in D$ such that $2 \leq |i - j| \leq k - 1$.*

We give the proof of this theorem in Section 4.1. It is geometric and uses the partition (2.1) of the set of all $(k-2)$ -subcubes of \mathcal{X} . We start by mapping each level set W_j : $j = 0, 1, \dots, k$ onto a point $P_j \in \mathbb{R}^2$. The mapping is built with the following transforms:

$$\begin{aligned} f_1 &: 2^{\mathcal{X}} \rightarrow \mathbb{R}^3, \\ f_1 &: W_j \mapsto (T_0^{(j)}, T_1^{(j)}, T_2^{(j)}) =: V_j \in \mathbb{R}^3, \\ f_2 &: \mathbb{R}^3 \rightarrow \mathbb{D}^2 \subset \mathbb{R}^3, \\ f_2 &: V_j \mapsto V_j / (T_0^{(j)} + T_1^{(j)} + T_2^{(j)}) =: Q_j, \end{aligned}$$

where the hyperplane $\mathbb{D}^2 := \{(v_1, v_2, v_3) \in \mathbb{R}^3 : v_1 + v_2 + v_3 = 1\}$. Note that within the non-negative orthant \mathbb{R}_+^3 this hyperplane covers one of the facets of the unit ball in the ℓ_1 -norm, i.e., $\mathbb{D}^2 \cap \mathbb{R}_+^3 = B_1(0, 1) \cap \mathbb{R}_+^3$. Next we introduce the transformation f_3 :

$$\begin{aligned} f_3 &: \mathbb{D}^2 \mapsto \mathbb{R}^2 \\ f_3 &: Q_j = (Q_j^{(1)}, Q_j^{(2)}, Q_j^{(3)}) \rightarrow (Q_j^{(1)}, Q_j^{(2)}) = P_j. \end{aligned}$$

Then we consider the closed polygonal chain $(P_0, P_1, \dots, P_k, P_0) \subset \mathbb{R}^2$. In order to prove that it bounds a convex polytope (polygon) in \mathbb{R}^2 , we need some technical lemmas to be proven.

Lemma 2.6. *Let $0 \leq j \leq k$. Then*

$$P_j^{(y)} = P_{k-j}^{(y)}.$$

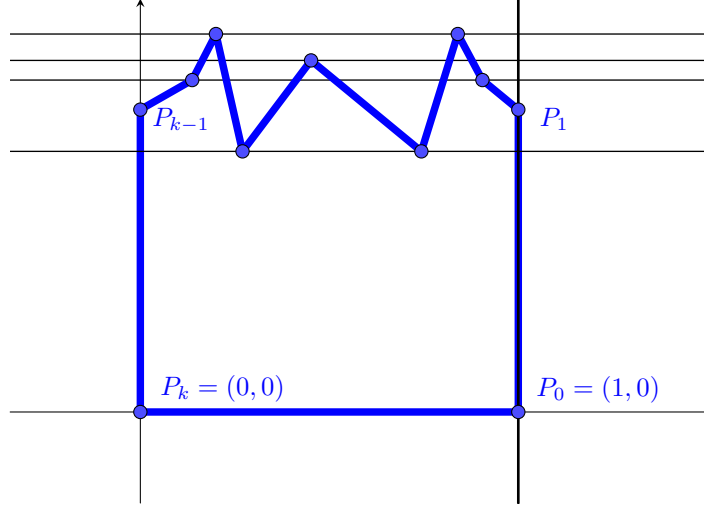


Figure 1: Illustration of the Lemma 2.6.

Now we state the following observation about the monotonicity of the slope of the points P_j .

Lemma 2.7. *Let $1 \leq j \leq k-2$. Then*

$$P_j^{(y)} / P_j^{(x)} = j / (k - j - 1).$$

Corollary 2.8. *Let $j \in \{0, 1, \dots, k-2\}$. Then the slope $P_j^{(y)} / P_j^{(x)}$ is an increasing function of j . Indeed, $P_0^{(y)} / P_0^{(x)} = 0/1 = 0 < P_1^{(y)} / P_1^{(x)}$. For larger values of j we apply Lemma 2.7.*

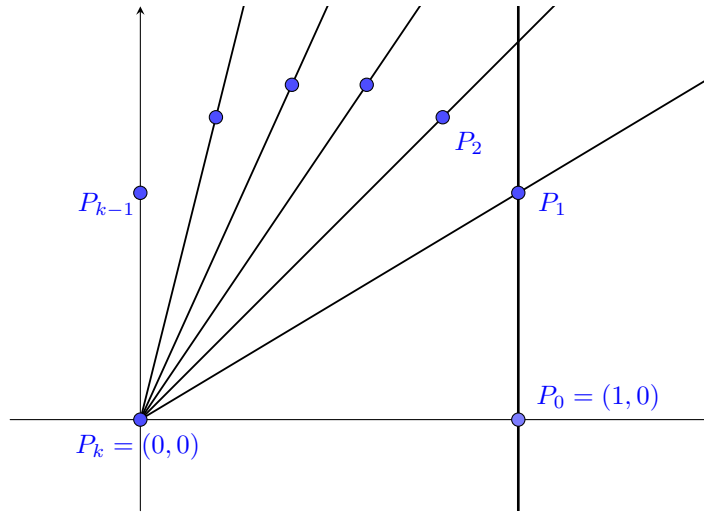


Figure 2: Illustration of the Corollary 2.8.

The rest of the vertices are $P_{k-1} = (0, \frac{1}{k-1})$ and $P_k = (0,0)$, cf. Figure 2.

Lemma 2.9. *For $j \leq \frac{k}{2}$ the value of $P_j^{(y)}$ increases with j . Moreover, for $j \geq \frac{k}{2} + 1$ it decreases with j .*

From Corollary 2.8 and Lemma 2.9 we deduce that the polygonal chain P is not self-intersecting. To prove that $(P_0, P_1, \dots, P_k, P_0)$ bounds a convex polytope, we show that no point P_j belongs to the interior of $\text{Conv}(\{P_0, P_1, \dots, P_j\})$. We prove that fact based on the connection between the level set W_D being of uniqueness for $(\mathcal{B}_2^k)_+$ and the relation between $\text{Conv}(\{Q_j : j \in D\})$ and $\text{int}(\text{Conv}(\{Q_0, Q_1, \dots, Q_k\}))$. For that reason, we use the following map ψ :

$$\begin{aligned} \psi : \mathbb{D}^2 &\rightarrow \mathbb{R} \\ \psi : Q_j &\mapsto f_2 \left(\sum_{x \in W_j} (T_0 + T_1 + T_2) \right) = f_2 \left(\sum_{x \in W_j} \varphi(x) \right). \end{aligned} \quad (2.3)$$

Remark 2.10. ψ is a convex function.

Lemma 2.11. Let $0 \leq j \leq k$. Then $\varphi(W_j) = \{0\}$ if and only if $\psi(Q_j) = 0$.

To finish the proof of convexity of $(P_0, P_1, \dots, P_k, P_0)$, we state and prove the following implications:

1. $(\forall j = 0, 1, \dots, k, W_j \text{ is not of uniqueness for } (\mathcal{B}_2^k)_+) \Rightarrow (\text{The chain } (P_0, P_1, \dots, P_k) \text{ is convex}),$
2. $(P \text{ is a convex polytope}) \Rightarrow (\forall j = 0, 1, \dots, k-1, W_{\{j, j+1\}} \text{ is not of uniqueness for } (\mathcal{B}_2^k)_+),$
3. $(W_{\{j, j+1\}} \text{ is not of uniqueness for } (\mathcal{B}_2^k)_+) \Rightarrow (W_j, W_{j+1} \text{ are not of uniqueness for } (\mathcal{B}_2^k)_+).$

Thus, for every $j = 0, \dots, k-1$ we obtain that

$$W_j \text{ is of uniqueness for } (\mathcal{B}_2^k)_+ \Leftrightarrow W_{j+1} \text{ is of uniqueness for } (\mathcal{B}_2^k)_+ \Leftrightarrow (P_0, P_1, \dots, P_k) \text{ is not convex},$$

which, with W_0 not being of uniqueness for $(\mathcal{B}_2^k)_+$, implies that (P_0, P_1, \dots, P_k) is convex.

Afterwards, we use the following lemma to show that if $2 \leq |i-j| \leq k-1$, then $W_{\{i,j\}}$ is of uniqueness for $(\mathcal{B}_2^k)_+$.

Lemma 2.12. Let j_1, j_2, j_3, j_4 be the vertices of $\text{Conv}(\{P_0, P_1, \dots, P_k\})$ and let $\varphi \in (\mathcal{B}_2^k)_+$. Assume that the segments (P_{j_1}, P_{j_3}) and (P_{j_2}, P_{j_4}) intersect. Then $\varphi(W_{\{j_1, j_3\}}) = \{0\}$ if and only if $\varphi(W_{\{j_2, j_4\}}) = \{0\}$.

The proof concludes with showing that the sets $W_{\{i, i+1\}}$ and $W_{\{0, k\}}$ are not of uniqueness for $(\mathcal{B}_2^k)_+$, which is done with an observation that $\text{Conv}(P_i, P_{i+1})$: $i = 0, 1, \dots, k-1$ and $\text{Conv}(P_0, P_k)$ are the edges of P . Analogously, for any other $\{i, j\}$ the set $W_{\{i, j\}}$ is of uniqueness for $(\mathcal{B}_2^k)_+$, since $\text{Conv}(P_i, P_j)$ is a proper diagonal of P .

3 Extremal properties

To improve the probability bounds to achieve the existence of the MLE for iid samples, we discuss the sets of uniqueness for $(\mathcal{B}_q^k)_+$ of the smallest cardinality. We start by introducing the following notation:

- $u(k, q) :=$ size of the smallest set of uniqueness for $(\mathcal{B}_q^k)_+$,
- $g(k, q) :=$ size of the smallest subset of $\mathcal{X} = \{-1, +1\}^k$ that intersects with every $(k-q)$ -subcube.

By Remark 2.1, we obtain the following result.

Corollary 3.1. For every k and q we have $g(k, q) \leq u(k, q)$.

The result above gives a lower bound for $u(k, q)$ that was studied as one of the Túrán-type problems, cf. [12]. The upper bounds for $u(k, q)$ can be established with the size of a known set of uniqueness either for $(\mathcal{B}_q^k)_+$ or for a larger class \mathcal{B}_q^k .

For larger values of q we use the fact proved by Graham et al. [12] that $g(k, q)$ can be bounded from below with the values of g for lower q .

Lemma 3.2. [12, Theorem 2. (ii)]

$$g(k, q) \geq 2g(k-1, q-1).$$

The exact value of $g(k, 2)$ was established by Kleitman and Spencer [15].

Lemma 3.3. [15, Theorem 1]

$$g(k, 2) = \min \left\{ r : \binom{r-1}{\lfloor k/2 \rfloor - 1} \geq k \right\}.$$

3.1 Partial results in general case

Below we state some partial results for the more general case $(\mathcal{B}_q^k)_+$.

To simplify the further notation, we denote the following equations and inequalities that refer to the level sets W_j , $j = 0, 1, \dots, k$:

$$0 = \sum_{x \in W_j} \varphi(x) = \dots = \binom{k-q}{j} T_0 + \dots + (k-q)T_{j-1} + T_j \quad (\text{equation } E_j)$$

$$0 \leq \sum_{x \in W_j} \varphi(x) = \dots = \binom{k-q}{j} T_0 + \dots + (k-q)T_{j-1} + T_j \quad (\text{inequality } I_j)$$

as E_j and I_j , respectively.

Theorem 3.4.

$$u(k, 3) \leq 2k.$$

Proof. In fact, we prove that the set $W_{\{1, k-1\}}$, which consists of k elements, is of uniqueness for $(\mathcal{B}_3^k)_+$. Indeed, the system $\{I_0, E_1, I_2, I_{k-2}, E_{k-1}, I_k\}$ is equivalent to

$$\begin{cases} T_0 & \geq 0 \\ (k-3)T_0 + T_1 & = 0 \\ \binom{k-3}{2}T_0 + (k-3)T_1 + T_2 & \geq 0 \\ T_1 + (k-3)T_2 + \binom{k-3}{2}T_3 & \geq 0 \\ T_2 + (k-3)T_3 & = 0 \\ T_3 & \geq 0 \end{cases}$$

This yields the following

$$\begin{aligned} 0 &\leq \binom{k-3}{2} T_0 + (k-3)T_1 + T_2 = \frac{(k-3)(k-4)}{2} T_0 + (k-3)T_1 + T_2 = \\ &= T_0 ((k-3)(k-4)/2 - (k-3)^2) + T_2 = T_0 \cdot (k-3)(k/2 - 2 - k + 3) + T_2 \leq T_0 \cdot (k-3)(1 - k/2) \leq 0, \end{aligned}$$

where the equality holds only if $T_2 = 0$. Therefore also $T_3 = 0$. Analogously, from (I_{k-2}) one can deduce that $T_1 = T_0 = 0$. Therefore $U = W_{\{1, k-1\}}$ is a set of uniqueness for $(\mathcal{B}_3^k)_+$. \square

Theorem 3.5. *For any $q \leq k$ there exists a set of uniqueness in $(\mathcal{B}_q^k)_+$ consisting of $O(k^{\lfloor q/2 \rfloor})$ elements.*

In fact, we prove a stronger result, i.e., that the set $U = W_{\{0, 1, 2, \dots, \lfloor q/2 \rfloor, k - \lfloor q/2 \rfloor, \dots, k-1, k\}}$, consisting of $2 \left(1 + k + \binom{k}{2} + \dots + \binom{k}{\lfloor q/2 \rfloor} \right)$ elements, is a set of uniqueness for the linear space \mathcal{B}_q^k . Therefore it is also of uniqueness for $(\mathcal{B}_q^k)_+$.

Proof. Let $\varphi(U) = \{0\}$. It implies the system of equations $\{E_0, E_1, \dots, E_{\lfloor \frac{q}{2} \rfloor}, E_{k - \lfloor \frac{q}{2} \rfloor}, \dots, E_{k-1}, E_k\}$, which, by Lemma 2.3, is the following:

$$\begin{cases} T_0 = 0 \\ (k-q)T_0 + T_1 = 0 \\ \binom{k-q}{2}T_0 + (k-q)T_1 + T_2 = 0 \\ \dots \\ \binom{k-q}{\lfloor q/2 \rfloor}T_0 + \binom{k-q}{\lfloor q/2 \rfloor - 1}T_1 + \dots + (k-q)T_{\lfloor q/2 \rfloor - 1} + T_{\lfloor q/2 \rfloor} = 0 \\ \text{-----} \\ T_{q - \lfloor q/2 \rfloor} + (k-q)T_{\lfloor q/2 \rfloor + 1} + \dots + \binom{k-q}{\lfloor q/2 \rfloor}T_q = 0 \\ \dots \\ T_{q-2} + (k-q)T_{q-1} + \binom{k-q}{2}T_q = 0 \\ T_{q-1} + (k-q)T_q = 0 \\ T_q = 0 \end{cases}$$

It is equivalent to the following matrix equation:

$$\left[\begin{array}{cccc|cccc} 1 & & & & & & & \\ k-q & 1 & & & & & & \\ \binom{k-q}{2} & k-q & 1 & & & & & \\ \vdots & \vdots & \vdots & \ddots & & & & \\ \binom{k-q}{\lfloor q/2 \rfloor} & \vdots & \vdots & \vdots & \vdots & & & 1 \end{array} \right] \cdot \left[\begin{array}{cccc} 1 & \dots & \dots & \dots & \binom{k-q}{\lfloor q/2 \rfloor} \\ & \ddots & \dots & \dots & \vdots \\ & & 1 & k-q & \binom{k-q}{2} \\ & & & 1 & k-q \\ & & & & 1 \end{array} \right] \cdot \begin{bmatrix} T_0 \\ T_1 \\ \vdots \\ T_q \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

The matrix on the left is an invertible $2(\lfloor q/2 \rfloor + 1) \times 2(\lfloor q/2 \rfloor + 1)$ matrix, therefore the only solution of the system of equation is $T_0 = T_1 = \dots = T_q = 0$ and thus $\varphi(\mathcal{X}) = \{0\}$. \square

We conclude the discussion on the smallest size of the set of uniqueness for $(\mathcal{B}_q^k)_+$ with the following theorem.

Theorem 3.6.

1. $u(k, 0) = 1$.
2. $u(k, 1) = 2$.
3. $\log k + \frac{1}{2} \log \log k + O(1) \leq u(k, 2) \leq k + 1$.
4. $2 \log(k-1) + \log \log(k-1) + O(1) \leq u(k, 3) \leq 2k$.
5. $u(k, k-1) = 2^{k-1}$.
6. $u(k, k) = 2^k$.

Moreover, for any $3 \leq q \leq k$ we obtain

$$2^{q-2} \log(k-q+2) + 2^{q-3} + \log \log(k-q+2) + O(1) \leq u(k, q) \leq O\left(k^{\lfloor q/2 \rfloor}\right).$$

Proof. (1):

When $q = 0$, all functions in $\mathcal{B} = (\mathcal{B}_0^k)_+$ are the constant positive functions. Therefore every non-empty subset of \mathcal{X} is of uniqueness, thus for every k the smallest size equals 1.

(2):

When $q = 1$, by Remark 2.1, the smallest set of uniqueness is $\{x, -x\}$, $x \in \mathcal{X}$, and its size equals 2.

(3):

By Lemma 2.4, the set $W_{\{1, k\}}$ is a minimal set of uniqueness for $(\mathcal{B}_2^k)_+$, therefore the smallest cardinality of the set of uniqueness does not exceed $k + 1$. The lower bound is the result of Lemma 3.3.

(4):

The upper bound for $u(k, 3)$ follows directly from Theorem 3.4. The lower bound is derived from Corollary 3.1, Lemma 3.2 and Lemma 3.3. Indeed,

$$u(k, 3) \stackrel{\text{Corollary 3.1}}{\geq} g(k, 3) \stackrel{\text{Lemma 3.2}}{\geq} 2g(k-1, 2) \stackrel{\text{Lemma 3.3}}{\geq} 2 \log(k-1) + \log \log(k-1) + O(1).$$

(5):

For every $k \geq 1$, the only minimal sets of uniqueness are $W_{\{0, 2, 4, \dots\}}$ and $W_{\{1, 3, 5, \dots\}}$, cf. Remark 2.1. Their sizes equal 2^{k-1} .

(6):

By Remark 2.1, the only set of uniqueness consists of 2^k elements.

(General case):

In general, the upper bound for $u(k, q)$ is obtained by Theorem 3.5. Similarly as in the part (4) of this proof, the lower bound comes from Corollary 3.1, Lemma 3.2 and Lemma 3.3. Indeed,

$$\begin{aligned} u(k, q) &\stackrel{\text{Corollary 3.1}}{\geq} g(k, q) \stackrel{\text{Lemma 3.2}}{\geq} 2g(k-1, q-1) \stackrel{\text{Lemma 3.2}}{\geq} \dots \stackrel{\text{Lemma 3.2}}{\geq} 2g(k-q+2, 2) \\ &\stackrel{\text{Lemma 3.3}}{\geq} 2^{q-2} \log(k-q+2) + 2^{q-3} \log \log(k-q+2) + O(1). \end{aligned}$$

\square

We finish this section with the following conjecture on the size of the smallest set of uniqueness for $(\mathcal{B}_2^k)_+$:

Conjecture 3.7. $u(k, 2) = k + 1$, i.e., for every k there are no sets of uniqueness having at most k elements.

4 Proofs

4.1 Main characterization

Proof of Theorem 2.5.

Step 1:

We commence the proof by mapping each single level set W_j : $j = 0, 1, \dots, k$ to a point $V_j := (T_0^{(j)}, T_1^{(j)}, T_2^{(j)}) \in \mathbb{R}^3$, where $T_i^{(j)}$ is the T_j from (2.2).

Next, we normalize each V_j to have its coordinates adding up to 1. Therefore, we perform the following transform:

$$Q_j := V_j / (T_0^{(j)} + T_1^{(j)} + T_2^{(j)}).$$

Observe that the set $\{Q_0, Q_1, \dots, Q_k\}$ lies in a two-dimensional hyperplane $\mathbb{D}^2 \subset \mathbb{R}^3$. Now we can project the set $\{Q_0, Q_1, \dots, Q_k\}$ onto \mathbb{R}^2 , where each $Q_j = (Q_j^{(x)}, Q_j^{(y)}, Q_j^{(z)})$ is mapped to $P_j := (Q_j^{(x)}, Q_j^{(y)})$.

Since projections preserve convexity, cf. [6, Sec. 2.3.2.], the polygonal chain $(Q_0, Q_1, \dots, Q_k, Q_0)$ bounds a convex polytope in \mathbb{R}^3 if and only if $(P_0, P_1, \dots, P_k, P_0)$ is a convex polytope in \mathbb{R}^2 .

Step 2:

Now we want to prove that $(P_0, P_1, \dots, P_k, P_0) \subset \mathbb{R}^2$ indeed bounds a convex polytope in \mathbb{R}^2 :

By Lemma 2.3, we obtain $P_0 = (1, 0)$ and $P_k = (0, 0)$. Moreover, according to Lemma 2.6, the second coordinates of P_j are distributed in a symmetric way, i.e. $P_j^{(y)} = P_{k-j}^{(y)}$.

From Corollary 2.8 we know that slope $P_j^{(y)} / P_j^{(x)}$ is an increasing function of j . Moreover, $(P_{k-1}^{(x)}, P_{k-1}^{(y)}) = (0, 1/(k-1))$ and $(P_k^{(y)}, P_k^{(y)}) = (0, 0)$. By Lemma 2.9 the value of $P_j^{(y)}$ increases with j for $j \leq k/2$ and decreases with j for $j \geq k/2 + 1$, see Figure 2.

Step 3:

Therefore, the polygonal chain $(P_0, P_1, \dots, P_k, P_0)$ is not self-intersecting. To prove that it is a convex polytope, we want to show that no point P_j is in the interior of $\text{Conv}(\{P_0, P_1, \dots, P_k\})$. Suppose that this statement is not true, i.e., there exists such j that $P_j \in \text{int}(\text{Conv}\{P_0, P_1, \dots, P_k\})$. Therefore also $Q_j \in \text{int}(\text{Conv}\{Q_0, Q_1, \dots, Q_k\})$, which implies the existence of such $\lambda_0, \lambda_1, \dots, \lambda_{j-1}, \lambda_{j+1}, \dots, \lambda_k > 0$ that $Q_j = \sum_{\substack{i=0 \\ i \neq j}}^k \lambda_i Q_i$. Thus $\psi(Q_j) = \sum_{\substack{i=0 \\ i \neq j}}^k \lambda_i \psi(Q_i)$ with ψ defined in (2.3).

Now assume that $\varphi(W_j) = \{0\}$. Then $\sum_{x \in W_j} \varphi(x) = 0$ and $\psi(Q_j) = f_2\left(\sum_{x \in W_j} \varphi(x)\right) = f_2(0) = 0$.

Thus also $\psi(Q_i) = 0$ for every $i \neq j$. But then $\varphi(\mathcal{X}) = \{0\}$, so W_j is a set of uniqueness for $(\mathcal{B}_2^k)_+$.

That way we obtain the following implication:

$$\left(\forall j = 0, 1, \dots, k, W_j \text{ is not of uniqueness for } (\mathcal{B}_2^k)_+ \right) \Rightarrow ((P_0, P_1, \dots, P_k, P_0) \text{ bounds a convex polytope}). \quad (4.1)$$

Now we proceed with proving the implication:

$$(P \text{ is a convex polytope}) \Rightarrow \left(\forall j = 0, 1, \dots, k-1, W_{\{j, j+1\}} \text{ is not of uniqueness for } (\mathcal{B}_2^k)_+ \right). \quad (4.2)$$

At first, we observe that $U = W_{\{0, k\}}$ is not of uniqueness for $(\mathcal{B}_2^k)_+$. To realize this, one may consider the function $\varphi = \mathbb{1}_{\{(+1, -1, *, *, \dots, *)\}} \in (\mathcal{B}_2^k)_+$, which equals one for every point of \mathcal{X} that has its first two coordinates fixed as $+1$ and -1 , respectively, and equals zero for every other point of \mathcal{X} . Note that φ vanishes both on $W_0 = \{(-1, -1, \dots, -1)\}$ and on $W_k = \{(+1, +1, \dots, +1)\}$. Now we proceed to show that for any $j \in \{0, 1, \dots, k-1\}$ the set $W_{\{j, j+1\}}$ is not of uniqueness for $(\mathcal{B}_2^k)_+$. For $j = 0$ it suffices to consider a function $\varphi = \mathbb{1}_{\{(+1, +1, *, *, \dots, *)\}} \in (\mathcal{B}_2^k)_+$ that vanishes on $U = W_{\{0, 1\}}$. An analogous argument holds for $j = k-1$ with a function $\varphi = \mathbb{1}_{\{(-1, -1, *, *, \dots, *)\}} \in (\mathcal{B}_2^k)_+$ vanishing on $U = W_{\{k-1, k\}}$. Now, let $1 \leq j \leq k-2$. Consider $U = W_{\{j, j+1\}}$ and suppose that $\varphi(U) = \{0\}$. As we assume, that P (and Q) is a convex polytope, we show the existence of such function ψ , that $\psi(Q_j) = \psi(Q_{j+1}) = 0$ and $\psi(Q_i) > 0$ for any $i \neq j, j+1$. To achieve this, we prove the existence of such T_0, T_1 and T_2 , that ensure

$\sum_{x \in W_j} \varphi(x) = \sum_{x \in W_{j+1}} \varphi(x) = 0$ and $\sum_{x \in W_i} \varphi(x) \geq 0$ for every $0 \leq i \leq k$. After establishing the correct T_0, T_1 and T_2 , we distribute them uniformly over all possible $(k-2)$ -subcubes of $\mathcal{X} = \{-1, +1\}^k$, which results in an exact function $\varphi \in (\mathcal{B}_2^k)_+$ vanishing on U , but not equal to zero on \mathcal{X} . In the following, we prove that such T_0, T_1 , and T_2 exist. Indeed, since Q is a convex polytope, there exists such vector $v \in \mathbb{R}^3$ that Q is supported by the hyperplanes $\{z : v^T z = v^T Q_j\}$ and $\{z : v^T z = v^T Q_{j+1}\}$, cf. [6, Sec. 2.5.2]. For $v \in \{Q_j\}^\perp \cap \{Q_{j+1}\}^\perp$ both hyperplanes go through Q_j and Q_{j+1} and their intersection is equal to $\text{aff}(\{Q_j, Q_{j+1}\})$. Moreover, $v^T z \geq 0$ for every $z \in Q$. Consider $(T_0, T_1, T_2) = v$. Notice that this vector results with a function $\varphi \in (\mathcal{B}_2^k)_+$ that vanishes on W_j and W_{j+1} , since $v^T Q_j = v^T Q_{j+1} = 0$. Similarly, $\varphi(W_j) > 0$, since $v^T Q_i > 0$ for $i \notin \{j, j+1\}$. Now, observe that for every $j = 0, 1, \dots, k-1$

$$\left(W_{\{j, j+1\}} \text{ is not of uniqueness for } (\mathcal{B}_2^k)_+\right) \Rightarrow \left(W_j, W_{j+1} \text{ are not of uniqueness for } (\mathcal{B}_2^k)_+\right). \quad (4.3)$$

From equations (4.1), (4.2) and (4.3), we obtain that for every $j = 0, 1, \dots, k-1$ the set W_j is of uniqueness for $(\mathcal{B}_2^k)_+$ if and only if W_{j+1} is of uniqueness for $(\mathcal{B}_2^k)_+$, too. But W_0 is not of uniqueness, neither is any other W_j . Moreover, it completes the proof that P is a convex polytope.

Step 4:

We complete the proof by showing that if $2 \leq j-i \leq k-1$, then $W_{\{i, j\}}$ is a set of uniqueness for $(\mathcal{B}_2^k)_+$. Without loss of generality, assume that $i < j$. Since P is a convex polytope, the segment (P_i, P_j) is one of its proper diagonals and intersects all diagonals $(P_{h_{in}}, P_{h_{ex}})$ for $i < h_{in} < j$ and $h_{ex} \in \{0, 1, \dots, i-1, j+1, \dots, k\}$. Therefore, by Lemma 2.12, $\varphi(W_{\{h_{in}, h_{ex}\}}) = \{0\}$. Then also $\varphi(\mathcal{X}) = \{0\}$, which implies that $W_{\{i, j\}}$ is a set of uniqueness for $(\mathcal{B}_2^k)_+$ and completes the proof. \square

4.2 Proofs of Lemmas

Proof of Lemma 2.6. By definition of P_j and Lemma 2.3, we have

$$P_j = \frac{\binom{k-2}{j}, \binom{k-2}{j-1}, \binom{k-2}{j-2}}{\binom{k-2}{j} + \binom{k-2}{j-1} + \binom{k-2}{j-2}}$$

and

$$P_{k-j} = \frac{\binom{k-2}{k-j}, \binom{k-2}{k-j-1}, \binom{k-2}{k-j-2}}{\binom{k-2}{k-j} + \binom{k-2}{k-j-1} + \binom{k-2}{k-j-2}} = \frac{\binom{k-2}{j-2}, \binom{k-2}{j-1}, \binom{k-2}{j}}{\binom{k-2}{j-2} + \binom{k-2}{j-1} + \binom{k-2}{j}}.$$

The claim is true as the denominator and the second coordinate of the numerator are the same. \square

Proof of Lemma 2.7. The proof is computational. By Lemma 2.3 we have

$$\begin{aligned} \frac{P_j^{(y)}}{P_j^{(x)}} &= \frac{\binom{k-2}{j-1}}{\binom{k-2}{j} + \binom{k-2}{j-1} + \binom{k-2}{j-2}} \cdot \frac{\binom{k-2}{j} + \binom{k-2}{j-1} + \binom{k-2}{j-2}}{\binom{k-2}{j}} = \frac{\binom{k-2}{j-1}}{\binom{k-2}{j}} \\ &= \frac{(k-2)!}{(j-1)!(k-j-1)!} \cdot \frac{j!(k-j-2)!}{(k-2)!} = \frac{j}{k-j-1}. \end{aligned}$$

\square

Proof of Lemma 2.9. Recall that $P_0^{(y)} = 0$ and $P_1^{(y)} = \frac{\binom{k-2}{0}}{\binom{k-2}{1} + \binom{k-2}{0} + \binom{k-2}{-1}} = \frac{1}{k-1} > 0$. For larger $j \leq \frac{k}{2}$ we will show that $P_{j+1}^{(y)}/P_j^{(y)}$ is greater than 1. Indeed, let $j \geq 1$. Then

$$\begin{aligned} \frac{P_{j+1}^{(y)}}{P_j^{(y)}} &= \frac{\binom{k-2}{j}}{\binom{k-2}{j+1} + \binom{k-2}{j} + \binom{k-2}{j-1}} \cdot \frac{\binom{k-2}{j} + \binom{k-2}{j-1} + \binom{k-2}{j-2}}{\binom{k-2}{j-1}} = \frac{k-j-1}{j} \cdot \frac{\binom{k-2}{j} + \binom{k-2}{j-1} + \binom{k-2}{j-2}}{\binom{k-2}{j+1} + \binom{k-2}{j} + \binom{k-2}{j-1}} \\ &= \frac{k-j-1}{j} \cdot \frac{\binom{k-2}{j-1} \cdot \left(\frac{k-j-1}{j} + 1 + \frac{j-1}{k-j}\right)}{\binom{k-2}{j-1} \cdot \left(\frac{(k-j-1)(k-j-2)}{j(j+1)} + \frac{k-j-1}{j} + 1\right)} = \frac{k-j-1}{j} \cdot \frac{\frac{(k-j)(k-j-1) + (k-j)j + (j-1)j}{j(k-j)}}{\frac{(k-j-1)(k-j-2) + (k-j-1)(j+1) + j(j+1)}{j(j+1)}} \\ &= \frac{k-j-1}{k-j} \cdot \frac{j+1}{j} \cdot \frac{(k-j)(k-j-1) + (k-j)j + (j-1)j}{(k-j-1)(k-j-2) + (k-j-1)(j+1) + j(j+1)} \\ &= \frac{k-j-1}{k-j} \cdot \frac{j+1}{j} \cdot \frac{(k-j)(k-1) + (j-1)j}{(k-j-1)(k-1) + j(j+1)} = (*). \end{aligned}$$

Expanding the last fraction allows to obtain

$$\begin{aligned} \frac{(k-j)(k-1) + (j-1)j}{(k-j-1)(k-1) + j(j+1)} &= \frac{(k-j)(k-1) + j(k-1) - j(k-j)}{(k-j-1)(k-1) + (j+1)(k-1) - (j+1)(k-j-1)} \\ &= \frac{k(k-1) - j(k-j)}{k(k-1) - (j+1)(k-j-1)} = 1 + \frac{(j+1)(k-j-1) - j(k-j)}{k(k-1) - (j+1)(k-j-1)} \\ &= 1 + \frac{(k-j) - (j-1)}{k(k-1) + (j+1)(k-j-1)} \end{aligned}$$

and therefore

$$(*) = \frac{1 + 1/j}{1 + 1/(k-j-1)} \cdot \left(1 + \frac{(k-j) - (j-1)}{k(k-1) + (j+1)(k-j-1)} \right) > \frac{1 + 1/j}{1 + 1/(k-j-1)} > 1.$$

The last two inequalities hold for $j \leq k/2$.

Analogously, for $j \geq k/2 + 1$, the value of $P_j^{(y)}$ decreases with j by Lemma 2.6. \square

Proof of Remark 2.10. Let $i, j \in \{0, 1, \dots, k\}$ and let $\lambda \in [0, 1]$. Then

$$\begin{aligned} \psi(\lambda Q_i + (1-\lambda)Q_j) &= f_2 \left(\lambda \sum_{x \in W_i} \varphi(x) + (1-\lambda) \sum_{x \in W_j} \varphi(x) \right) \stackrel{(*)}{=} \lambda f_2 \left(\sum_{x \in W_i} \varphi(x) \right) + (1-\lambda) f_2 \left(\sum_{x \in W_j} \varphi(x) \right) \\ &= \lambda \psi(Q_i) + (1-\lambda) \psi(Q_j). \end{aligned}$$

The equation with the asterisk follows from the convexity of the map f_2 on \mathbb{R}_+^3 , cf. [6, Sec. 2.3.2.]. Therefore ψ is also a convex map. \square

Proof of Lemma 2.11.

$$\varphi(W_j) = \{0\} \Leftrightarrow \sum_{x \in W_j} \varphi(x) = 0 \Leftrightarrow f_2 \left(\sum_{x \in W_j} \varphi(x) \right) = 0 = \psi(Q_j).$$

\square

Proof of Lemma 2.12. Assume that the segments (P_{j_1}, P_{j_3}) and (P_{j_2}, P_{j_4}) intersect. Then also do the segments (Q_{j_1}, Q_{j_3}) and (Q_{j_2}, Q_{j_4}) , because the map f_3^{-1} preserves convexity. But it means that there exist such $\lambda, \mu \in (0, 1)$ that $\lambda Q_{j_1} + (1-\lambda)Q_{j_3} = \mu Q_{j_2} + (1-\mu)Q_{j_4}$. Therefore

$$\begin{aligned} \varphi(W_{\{j_1, j_3\}}) = \{0\} &\Leftrightarrow f_2 \left(\sum_{x \in W_{\{j_1, j_3\}}} \varphi(x) \right) = 0 \Leftrightarrow \psi(Q_{j_1}) = \psi(Q_{j_3}) = 0 \Leftrightarrow \lambda \psi(Q_{j_1}) + (1-\lambda) \psi(Q_{j_3}) = 0 \\ &\Leftrightarrow \psi(\lambda Q_{j_1} + (1-\lambda)Q_{j_3}) = 0 = \psi(\mu Q_{j_2} + (1-\mu)Q_{j_4}) \Leftrightarrow \psi(Q_{j_2}) = \psi(Q_{j_4}) = 0 \Leftrightarrow \varphi(W_{\{j_2, j_4\}}) = \{0\}. \end{aligned}$$

\square

Acknowledgements

We thank Krzysztof Bogdan and Piotr Zwiernik for their insightful comments.

References

- [1] N. Alon, A. Krech, and T. Szabó. Turán's theorem in the hypercube. *SIAM J. Discrete Math.*, 21(1):66–72, 2007.
- [2] A. Baber. Some results in extremal combinatorics. *PhD Thesis*, 3, 2011.
- [3] O. Barndorff-Nielsen. *Information and exponential families in statistical theory*. John Wiley & Sons Ltd., Chichester, 1978. Wiley Series in Probability and Mathematical Statistics.
- [4] K. Bogdan and M. Bogdan. On existence of maximum likelihood estimators in exponential families. *Statistics*, 34(2):137–149, 2000.

- [5] K. Bogdan, M. Bosy, and T. Skalski. Maximum likelihood estimation for discrete exponential families and random graphs. *ALEA Lat. Am. J. Probab. Math. Stat.*, 19(1):1045–1070, 2022.
- [6] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, USA, 2004.
- [7] S. G. Brush. History of the Lenz-Ising model. *Rev. Mod. Phys.*, 39:883–893, Oct 1967.
- [8] B. R. Crain. Estimation of distributions using orthogonal expansions. *Ann. Statist.*, 2:454–463, 1974. Collection of articles dedicated to Jerzy Neyman on his 80th birthday.
- [9] B. R. Crain. Exponential models, maximum likelihood estimation, and the Haar condition. *J. Amer. Statist. Assoc.*, 71(355):737–740, 1976.
- [10] A. Dembo and A. Montanari. Ising models on locally tree-like graphs. *Ann. Appl. Probab.*, 20(2):565–592, 2010.
- [11] S. Dutt and J. P. Hayes. On allocating subcubes in a hypercube multiprocessor. In *Proceedings of the Third Conference on Hypercube Concurrent Computers and Applications: Architecture, Software, Computer Systems, and General Issues - Volume 1*, C3P, pages 801–810, New York, NY, USA, 1988. Association for Computing Machinery.
- [12] N. Graham, F. Harary, M. Livingston, and Q. F. Stout. Subcube fault-tolerance in hypercubes. *Inform. and Comput.*, 102(2):280–314, 1993.
- [13] J. Jendrej, K. Oleszkiewicz, and J. O. Wojtaszczyk. On some extensions of the FKN theorem. *Theory Comput.*, 11:445–469, 2015.
- [14] J. K. Johnson, D. Oyen, M. Chertkov, and P. Netrapalli. Learning planar Ising models. *J. Mach. Learn. Res.*, 17:Paper No. 215, 26, 2016.
- [15] D. J. Kleitman and J. Spencer. Families of k -independent sets. *Discrete Math.*, 6:255–262, 1973.
- [16] S. L. Lauritzen. Extreme point models in statistics. *Scand. J. Statist.*, 11(2):65–91, 1984.
- [17] E. L. Lehmann and G. Casella. *Theory of point estimation*. Springer Texts in Statistics. Springer-Verlag, New York, second edition, 1998.
- [18] M. J. Wainwright and M. I. Jordan. *Graphical Models, Exponential Families, and Variational Inference*, volume 1. 2008.
- [19] N. Wang, J. Rauh, and H. Massam. Approximating faces of marginal polytopes in discrete hierarchical models. *Ann. Statist.*, 47(3):1203–1233, 2019.