SETS OF UNIQUENESS FOR THE ISING MODEL: $u(k,2) = k+1$, AND THE ORDER OF $u(k,q)$

KOREA SUPERINTELLIGENCE LABS

Abstract. On the discrete cube $\mathcal{X} = \{-1,+1\}^k$ let B_q^k be the span of the Walsh functions of degree at most q and let $(B_q^k)_+$ be its cone of non-negative members; B_2^k is the Ising model on the complete graph. A set $U \subseteq \mathcal{X}$ is a *set of uniqueness* for $(B_q^k)_+$ if 0 is the only non-negative $\varphi \in B_q^k$ vanishing on U — the condition, due to Bogdan, Bosy and Skalski, under which the maximum likelihood estimator for the corresponding exponential family and a sample supported on U exists — and $u(k,q)$ denotes the least size of such a set. Skalski and Stroiński proved $u(k,2) \le k+1$, proved lower bounds of order $\log k$, and conjectured that $u(k,2) = k+1$. We prove the conjecture in the form $u(k,2) = k+1$ for every $k \ge 3$, and observe that it is false as literally stated at $k = 2$, where $u(2,2) = 4$. The lower bound is a special case of a general one: every set of uniqueness for $(B_{2m}^k)_+$ has at least $\dim B_m^k = \sum_{i \le m} \binom{k}{i}$ points, so that $u(k,q) \ge \sum_{i \le \lfloor q/2 \rfloor} \binom{k}{i}$ for every q. Combined with the upper bound $u(k,q) = O(k^{\lfloor q/2 \rfloor})$ of Skalski and Stroiński this determines the order of $u(k,q)$ for each fixed q, namely $u(k,q) = \Theta(k^{\lfloor q/2 \rfloor})$, where the previously known lower bounds were logarithmic in k; and it is attained at $q = k-1$ for odd k. The mechanism of the lower bound is the classical one of Rao's bound for orthogonal arrays and of lower bounds for positive cubature formulae, transplanted to the weaker hypothesis; we say so plainly, and the statement, not the technique, is what is offered as new. We also correct a transcription of a formula of Kleitman and Spencer quoted in the source. All numbered results are machine-checked in Lean 4 against Mathlib and are general in k; Section 7 says exactly what is formalized, what is quoted, and which few numbers are computations rather than theorems.

1. Introduction

The objects. Fix $k \in \mathbb{N}$ and let $\mathcal{X} = \{-1,+1\}^k$ be the discrete cube with the uniform weight $\mu(x) = 2^{-k}$. For $j = 1, \dots, k$ the *Rademacher functions* are $r_j(x) = x_j$, with $r_0 \equiv 1$, and for $L \subseteq \{1, \dots, k\}$ the *Walsh function* w_L is

$$w_L(x) \;=\; \prod_{j \in L} r_j(x), \qquad w_\emptyset \equiv 1.$$

For $0 \le q \le k$ put

$$B_q^k \;=\; \operatorname{Lin}\big\{\, w_L \;:\; L \subseteq \{1,\dots,k\},\ |L| \le q \,\big\}, \qquad (B_q^k)_+ \;=\; \{\, \varphi \in B_q^k \;:\; \varphi \ge 0 \text{ pointwise} \,\}.$$

The class B_2^k is the Ising model on the complete graph: the exponential family generated by it consists of the Boltzmann measures $\mu(x) \propto \exp\big(B \sum_i x_i + \beta \sum_{i<j} x_i x_j\big)$.

Bogdan, Bosy and Skalski [2] characterised the existence of the maximum likelihood estimator in such a family by the following notion, which is the subject of this paper. We quote the definition as it is printed in [1], writing B and φ for its $\mathcal{B}$ and ϕ.

> **Definition.** Let $U \subset \mathcal{X}$. Let B be a fixed linear subspace of $\mathbb{R}^{\mathcal{X}}$. We say that U is a set of uniqueness for B if $\varphi = 0$ is the only function in B such that $\varphi = 0$ on U. Similarly, we say that U is a *set of uniqueness* for B_+ if $\varphi = 0$ is the only function in B_+ such that $\varphi = 0$ on U.

The maximum likelihood estimator for the family generated by B and a sample $x_1, \dots, x_n$ exists if and only if $\{x_1, \dots, x_n\}$ is a set of uniqueness for B_+ (Theorem 2.2 of [2]); so the smallest sets of

uniqueness are the smallest sample supports from which the model can be estimated at all. Following [1] we write

$$u(k,q) \;=\; \text{the size of the smallest set of uniqueness for } (B_q^k)_+,$$

and, for the comparison quantity used there,

$$g(k,q) \;=\; \text{the size of the smallest subset of } \mathcal{X} \text{ that intersects every } (k-q)\text{-subcube.}$$

Finally, W_D denotes the level set $\{x \in \mathcal{X} : \operatorname{dist}(x, \perp) \in D\}$, where $\perp = (-1, \dots, -1)$ and dist is Hamming distance; we abbreviate $W_{\{a,b\}}$ to $W_{a,b}$, as [1] does.

The source and its questions. Skalski and Stroiński [1] compute $u(k,q)$ among level sets. Their Theorem 2.5 characterises which W_D are sets of uniqueness for the Ising class (W_D works if and only if some $i, j \in D$ satisfy $2 \le |i-j| \le k-1$); their Lemma 2.4 exhibits $W_{1,k}$, of size $k+1$, and so gives $u(k,2) \le k+1$; their Theorem 3.5 gives $u(k,q) = O(k^{\lfloor q/2 \rfloor})$ for every $q \le k$; and their Theorem 3.6 collects the values

$$u(k,0) = 1, \quad u(k,1) = 2, \quad u(k,k-1) = 2^{k-1}, \quad u(k,k) = 2^k$$

together with lower bounds. Every lower bound there passes through $g(k,q)$, a Turán-type covering number governed for $q = 2$ by a theorem of Kleitman and Spencer [3], and every one of them is *logarithmic* in k: for $q = 2$ the source prints $\log k + \frac{1}{2}\log\log k + O(1) \le u(k,2) \le k+1$, and for $3 \le q \le k$ it prints

$$2^{q-2}\log(k-q+2) + 2^{q-3} + \log\log(k-q+2) + O(1) \;\le\; u(k,q) \;\le\; O\big(k^{\lfloor q/2 \rfloor}\big).$$

The paper ends with the following conjecture, which we quote as printed.

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

So two gaps are left open: the exponential gap between $\Theta(\log k)$ and $k+1$ at $q = 2$, and the gap between a logarithmic lower bound and a polynomial upper bound for every larger fixed q.

What is settled here.

- *The conjecture* (Theorem 4.2). $u(k,2) = k+1$ for every $k \ge 3$: $k+1$ is the least cardinality of a set of uniqueness for the Ising class. The upper half is the source's Lemma 2.4, reproved in Section 4 by an explicit positive weight (Proposition 4.1); the lower half, Theorem 3.2, is the half that was open.
- *A general lower bound* (Theorem 3.1). Every set of uniqueness for $(B_{2m}^k)_+$ has at least $\dim B_m^k = \sum_{i \le m} \binom{k}{i}$ points; that is, $u(k,2m) \ge \sum_{i \le m} \binom{k}{i}$. At $m = 1$ this is $u(k,2) \ge k+1$.
- *The order of* $u(k,q)$ (Theorem 3.3 and Corollary 3.4). For every q, $u(k,q) \ge \sum_{i \le \lfloor q/2 \rfloor} \binom{k}{i}$. With the source's Theorem 3.5 this gives $u(k,q) = \Theta(k^{\lfloor q/2 \rfloor})$ for each fixed q, in place of lower bounds of order $\log k$. The bound is attained at $q = k-1$ for odd k (Remark 3.5), so it cannot be improved in general.
- *Two corrections to* [1] (Section 6). The conjecture is false at $k = 2$ as literally stated, since $u(2,2) = 4 > 3$ (Proposition 6.1); the correct hypothesis is $k \ge 3$, and the conjecture's second clause — that no set of at most k elements is a set of uniqueness — is true for every k. And the formula of Kleitman and Spencer quoted in the source's Lemma 3.3 prints $\lfloor k/2 \rfloor$ where it must read $\lfloor r/2 \rfloor$ (Proposition 6.2).

Section 5 reproves, by the same certificate device, the values $u(k,0) = 1$, $u(k,1) = 2$ and $u(k,k) = 2^k$ of the source's Theorem 3.6. Those are not new; they are included because they are the controls that keep the new statements honest. The value $u(k,1) = 2$ in particular shows that the hypothesis $q = 2$ in Theorem 3.2 is not decorative: for the Rademacher class B_1^k the answer is 2 and not $k+1$, so the step that squares an affine function — the one place where the degree goes up — is load-bearing. Similarly $u(k,k) = 2^k$ shows that no bound $u(k,q) \le \mathrm{poly}(k)$ can hold uniformly in

q, and Proposition 5.4 exhibits a second level set of $k+1$ points that is *not* a set of uniqueness, so that Theorem 4.2 is a statement about which $k+1$ points and not about the number alone.

Where the new content begins, and where the argument comes from. Two things should be said plainly before the proofs.

First, on the mechanism. The proof of Theorem 3.1 is three lines: if $|U| < \dim B_m^k$ then some nonzero $\ell \in B_m^k$ vanishes on U, and ℓ^2 is a non-negative, not identically zero member of B_{2m}^k vanishing on U. This is the classical lower-bound argument for positive cubature formulae, where it shows that a formula exact to degree $2m$ needs at least $\dim P_m$ nodes [5, 6], and it is the argument behind Rao's bound for orthogonal arrays [4], which for a binary array of strength $2m$ on k factors reads $N \geq \sum_{i \leq m} \binom{k}{i}$ — the same cube, the same Walsh basis, and numerically the same bound. Rao's bound does not imply Theorem 3.1: an orthogonal array of strength $2m$ is a set on which every Walsh function of degree between 1 and $2m$ averages to zero, which is a far stronger hypothesis on U than being a set of uniqueness. So what is offered as new here is the *statement* — neither [1] nor [2] contains any bound of this shape, and both leave the gap open — and not the technique, which a referee may fairly call textbook. Section 7 records the searches behind that claim of novelty and the residual risk in it.

Second, on the range of k. The source's Corollary 3.1 gives $g(k,2) \leq u(k,2)$, and $g(k,2) = k+1$ for $k = 3, 4, 5$ (a finite computation, reported in Remark 6.3). For those three values of k the conjecture therefore already follows from the source's own tools together with that computation, and $g(3,2) = 4$, $g(4,2) = 5$, $g(5,2) = 6$ are the last values for which it does: $g(6,2) = 6 < 7$, and from there $g(k,2)$ falls away logarithmically. So the new content of Theorems 3.2 and 4.2 begins at $k = 6$. We state them for all k because the proof is uniform in k and because the source states the conjecture for all k.

What is not claimed. We prove no upper bound beyond $u(k,2) \leq k+1$. In particular the upper half of Corollary 3.4 is quoted from [1], not reproved and not formalized, and the value $u(k,k-1) = 2^{k-1}$ used in Remark 3.5 is quoted from [1] as well. We do not determine $u(k,q)$ exactly for $3 \leq q \leq k-2$; for $q = 3$ our lower bound $k+1$ and the source's upper bound $2k$ leave a factor of 2. We do not decide whether the bound of Theorem 3.1 is attained for $2m < k-1$. We do not reprove the source's Theorem 2.5 or its Theorem 3.5, and nothing here bears on the asymptotics of $g(k,q)$ beyond the transcription corrected in Proposition 6.2.

2. Preliminaries

Throughout, k and q are non-negative integers, $\mathcal{X} = \{-1,+1\}^k$, and functions on $\mathcal{X}$ are real-valued. We write $\perp = (-1,\dots,-1)$ and $\top = (+1,\dots,+1)$ for the two antipodal points, u_j for the point whose only $+1$ coordinate is the j-th, and δ_y for the indicator of $y \in \mathcal{X}$. Sums $\sum_x$ run over $\mathcal{X}$. Binomial coefficients $\binom{k}{i}$ are 0 for $i > k$, so that $\dim B_m^k = \sum_{i=0}^m \binom{k}{i}$ holds without a caveat for $m > k$ as well.

The elementary Walsh calculus we use is the following.

Lemma 2.1. *For all $L, M \subseteq \{1,\dots,k\}$ and $x \in \mathcal{X}$:*

1. $w_L(x)\, w_M(x) = w_{L \triangle M}(x)$, *where $\triangle$ is symmetric difference;*
2. $\sum_x w_L(x) = 2^k$ *if $L = \emptyset$ and 0 otherwise, whence* $\sum_x w_L(x) w_M(x) = 2^k \cdot [L = M]$;
3. *consequently* $\{w_L\}_{|L| \leq q}$ *is linearly independent and* $\dim B_q^k = \sum_{i \leq q} \binom{k}{i}$;
4. $B_p^k \cdot B_q^k \subseteq B_{p+q}^k$; *in particular $\ell \in B_m^k$ implies $\ell^2 \in B_{2m}^k$;*
5. $w_L(\perp) = (-1)^{|L|}$, $w_L(\top) = 1$, *and* $\sum_{j=1}^k w_L(u_j) = (k - 2|L|)(-1)^{|L|}$.

Proof. (1) Since $x_j^2 = 1$ on the cube, $\prod_{j \in L} x_j \cdot \prod_{j \in M} x_j = \prod_{j \in L \triangle M} x_j \cdot \prod_{j \in L \cap M} x_j^2 = w_{L \triangle M}(x)$; this is the only place the cube, as opposed to a general product space, is used. (2) For $L \neq \emptyset$ pick $j \in L$; flipping the j-th coordinate is an involution of $\mathcal{X}$ that negates w_L, so the sum equals its own negative. The second identity follows from (1). (3) Immediate from (2), since a vanishing combination paired

against w_M returns 2^k times its M-th coefficient; the dimension count is the number of L with $|L| \le q$. (4) By (1), $w_L w_M = w_{L \triangle M}$ and $|L \triangle M| \le |L| + |M|$. (5) Direct computation: $w_L(u_j) = -(-1)^{|L|}$ if $j \in L$ and $(-1)^{|L|}$ otherwise, so the sum over j is $\big((k - |L|) - |L|\big)(-1)^{|L|}$. □

We use throughout, without further comment, that $B_p^k \subseteq B_q^k$ for $p \le q$, hence that a set of uniqueness for $(B_q^k)_+$ is a set of uniqueness for $(B_p^k)_+$ and that $q \mapsto u(k,q)$ is non-decreasing; this is Remark 1.3 of [1], which attributes it to Remark 5.6 of [2].

The upper bounds in this paper all come from one elementary observation, the dual side of the definition.

Lemma 2.2 (certificate criterion)**.** *Let $B \subseteq \mathbb{R}^{\mathcal{X}}$ be a linear subspace and $U \subseteq \mathcal{X}$. Suppose there is $c : \mathcal{X} \to \mathbb{R}$ with $c(x) > 0$ for every $x \notin U$ and $\sum_x c(x)\varphi(x) = 0$ for every $\varphi \in B$. Then U is a set of uniqueness for B_+.*

Proof. Let $\varphi \in B_+$ vanish on U. Then $0 = \sum_x c(x)\varphi(x) = \sum_{x \notin U} c(x)\varphi(x)$ is a sum of non-negative terms, so each vanishes, and $c(x) > 0$ off U forces $\varphi = 0$ off U as well. □

It suffices to test the condition on the generators w_L, $|L| \le q$, since both sides are linear in φ. Lemma 2.2 is not new and we claim nothing for it: it is the easy direction of a theorem of the alternative. Its converse holds as well — if U is a set of uniqueness for B_+ then such a c exists, by applying Stiemke's lemma [7] to $V = \{\varphi \in B : \varphi|_U = 0\}$ and correcting the resulting strictly positive vector on U, where it is unconstrained — and §2.2 of [2] already writes the primal program of which this is the dual. We use only the direction proved above, and only that direction is formalized.

3. The lower bound

Theorem 3.1. *Let $k, m \ge 0$ and let $U \subseteq \mathcal{X}$ with $|U| < \dim B_m^k = \sum_{i=0}^m \binom{k}{i}$. Then U is not a set of uniqueness for $(B_{2m}^k)_+$. Equivalently, every set of uniqueness for $(B_{2m}^k)_+$ has at least $\sum_{i \le m} \binom{k}{i}$ elements:*

$$u(k, 2m) \;\ge\; \sum_{i=0}^{m} \binom{k}{i}.$$

Proof. Consider the evaluation map $B_m^k \to \mathbb{R}^U$, $\ell \mapsto \ell|_U$. Its source has dimension $\sum_{i \le m} \binom{k}{i}$ by Lemma 2.1(3) and its target has dimension $|U|$, which is smaller, so by rank–nullity there is $\ell \in B_m^k$, $\ell \ne 0$, with $\ell|_U = 0$. (Concretely: the $|U|$ homogeneous linear conditions $\ell(u) = 0$, $u \in U$, on the coefficient vector of ℓ have a nonzero solution, and a nonzero coefficient vector gives a nonzero function by the linear independence of the Walsh system.) Put $\varphi = \ell^2$. Then $\varphi \ge 0$ pointwise; φ vanishes on U; φ is not identically zero, since ℓ is not; and $\varphi \in B_{2m}^k$ by Lemma 2.1(4), which is where $x_j^2 = 1$ is used. So φ is a nonzero member of $(B_{2m}^k)_+$ vanishing on U, and U is not a set of uniqueness. □

Theorem 3.2. *For every k and every $U \subseteq \mathcal{X}$ with $|U| \le k$, the set U is not a set of uniqueness for the Ising class $(B_2^k)_+$. Hence $u(k, 2) \ge k + 1$.*

Proof. This is the case $m = 1$ of Theorem 3.1: $\dim B_1^k = k + 1$, so $|U| \le k$ is the hypothesis $|U| < \dim B_1^k$. Unwound, the argument is the one described in Section 1: the k or fewer conditions $\ell(u) = 0$ on the $k + 1$ coefficients of an affine $\ell = a_0 + \sum_j a_j x_j$ have a nonzero solution, and ℓ^2 lies in B_2^k because $x_j^2 = 1$, is non-negative, is not identically zero, and vanishes on U. □

Theorem 3.2 is exactly the second clause of the conjecture quoted in Section 1 — "for every k there are no sets of uniqueness having at most k elements" — and it holds for every k without exception, including $k = 2$. The first clause needs $k \ge 3$; see Section 6.

Theorem 3.3. *For all k and q, every set of uniqueness for $(B_q^k)_+$ has at least $\sum_{i\le\lfloor q/2\rfloor}\binom{k}{i}$ elements:*

$$u(k,q)\ \ge\ \sum_{i=0}^{\lfloor q/2\rfloor}\binom{k}{i}.$$

Proof. Put $m=\lfloor q/2\rfloor$, so $2m\le q$. A set of uniqueness for $(B_q^k)_+$ is one for $(B_{2m}^k)_+$ by monotonicity, and Theorem 3.1 applies. □

Corollary 3.4. *For each fixed $q\ge 0$, $u(k,q)=\Theta\big(k^{\lfloor q/2\rfloor}\big)$ as $k\to\infty$.*

Proof. The lower bound is Theorem 3.3, since $\sum_{i\le m}\binom{k}{i}\ge\binom{k}{m}\ge(k-m)^m/m!$ for $m=\lfloor q/2\rfloor$ fixed. The upper bound $u(k,q)=O(k^{\lfloor q/2\rfloor})$ is Theorem 3.5 of [1], quoted and not reproved here. □

Theorem 3.5 of [1] reads, verbatim, "For any $q\le k$ there exists a set of uniqueness in $(B_q^k)_+$ consisting of $O(k^{\lfloor q/2\rfloor})$ elements", and the sentence following it states the stronger explicit form that its proof establishes: the level set $W_{\{0,1,\dots,\lfloor q/2\rfloor\}\cup\{k-\lfloor q/2\rfloor,\dots,k\}}$, of $2\big(1+k+\binom{k}{2}+\cdots+\binom{k}{\lfloor q/2\rfloor}\big)$ points, is a set of uniqueness for the linear space B_q^k and hence for its cone. In that explicit form the bound is uniform in q, and combining it with Theorem 3.3 gives

$$\sum_{i\le\lfloor q/2\rfloor}\binom{k}{i}\ \le\ u(k,q)\ \le\ 2\sum_{i\le\lfloor q/2\rfloor}\binom{k}{i}\qquad(q\le k),$$

the lower half proved here and the upper half quoted. Corollary 3.4 is stated for fixed q, which is all that the $\Theta(k^{\lfloor q/2\rfloor})$ form needs.

Corollary 3.4 replaces, for every fixed q, the lower bounds of order $\log k$ recorded in [1]; at $q=2$ it is the difference between $\Theta(\log k)$ and $k+1$.

Remark 3.5. The bound of Theorem 3.1 cannot be improved in general. Take $q=k-1$ with k odd; then $\lfloor q/2\rfloor=(k-1)/2$ and $\sum_{i\le(k-1)/2}\binom{k}{i}=2^{k-1}$ by the symmetry of the binomial coefficients, while $u(k,k-1)=2^{k-1}$ by Theorem 3.6(5) of [1]. So Theorem 3.3 is an equality there. We checked the case $k=5$ independently: the level set $W_{\{0,2,4\}}$ of the 16 points of even weight is a set of uniqueness for $(B_4^5)_+$, and $1+5+10=16$. Whether the bound is attained for $2m<k-1$ we do not know; the first open case is $q=4$, where $1+k+\binom{k}{2}\le u(k,4)\le 2\big(1+k+\binom{k}{2}\big)$, the upper bound being the explicit form of Theorem 3.5 of [1] quoted after Corollary 3.4.

4. The upper bound and the conjecture

Proposition 4.1. *For $k\ge 2$ the level set $W_{1,k}=\{u_1,\dots,u_k,\top\}$ has $k+1$ elements, and for $k\ge 3$ it is a set of uniqueness for $(B_2^k)_+$. Hence $u(k,2)\le k+1$ for $k\ge 3$.*

Proof. The $k+1$ listed points are distinct for $k\ge 2$. For the second assertion we apply Lemma 2.2 with the weight

$$c\ =\ 1+2^{k-2}\Big((k-3)\,\delta_\perp-\sum_{j=1}^{k}\delta_{u_j}-\delta_\top\Big),$$

that is, $c(\perp)=1+(k-3)2^{k-2}$, $c(u_j)=c(\top)=1-2^{k-2}$, and $c(x)=1$ for every other x. Positivity off $W_{1,k}$ holds because the only point outside $W_{1,k}$ at which $c\ne 1$ is $\perp$, where $c(\perp)=1+(k-3)2^{k-2}>0$ for $k\ge 3$. For the annihilation, let $|L|=t\le 2$ and substitute the values of Lemma 2.1(2),(5):

$$\sum_x c(x)w_L(x)=2^k[t=0]+2^{k-2}\Big((k-3)(-1)^t-(k-2t)(-1)^t-1\Big),$$

which for $t=0$ is $2^k+2^{k-2}\big((k-3)-k-1\big)=2^k-4\cdot 2^{k-2}=0$; for $t=1$ is $2^{k-2}\big(-(k-3)+(k-2)-1\big)=0$; and for $t=2$ is $2^{k-2}\big((k-3)-(k-4)-1\big)=0$. □

Proposition 4.1 is Lemma 2.4 of [1] and is not new; the source proves it by a decomposition of the cube into subcubes. The proof above is included because the explicit weight c is the object that the source's argument leaves implicit, because the same device gives the source's other values in two lines each (Section 5), and because it is short enough to formalize.

Theorem 4.2. *Let $k \geq 3$. Then $k+1$ is the least cardinality of a set of uniqueness for the Ising class $(B_2^k)_+$; that is,*

$$u(k,2) \;=\; k+1.$$

Proof. The set $W_{1,k}$ of Proposition 4.1 is a set of uniqueness with $k+1$ elements, so $k+1$ belongs to the set of cardinalities in question. By Theorem 3.2 no set of uniqueness for $(B_2^k)_+$ has k or fewer elements, so $k+1$ is a lower bound for that set. Hence it is its least element, and the infimum defining $u(k,2)$ equals $k+1$. □

This is the conjecture of [1], with the hypothesis $k \geq 3$ that Proposition 6.1 below shows to be necessary. As explained in Section 1, the cases $k = 3, 4, 5$ were already within reach of the source's own tools; the content of Theorem 4.2 is $k \geq 6$.

5. Controls, and the source's other values

The four statements of this section are not new — three are printed in [1] and the fourth is an instance of its Theorem 2.5 — but each of them constrains how Theorems 3.1, 3.2 and 4.2 may be read, and each is proved here from the definitions by the device of Lemma 2.2.

Proposition 5.1. *$u(k,0) = 1$ for every k.*

Proof. B_0^k consists of the constants, so every nonempty U is a set of uniqueness and the empty set is not (the constant 1 vanishes on it vacuously and is a nonzero non-negative member). □

Proposition 5.2. *$u(k,1) = 2$ for every $k \geq 1$, with witness $W_{0,k} = \{\bot, \top\}$.*

Proof. For the upper bound apply Lemma 2.2 with $c = 1 - 2^{k-1}(\delta_\bot + \delta_\top)$, which is 1 off $\{\bot, \top\}$ and satisfies $\sum_x c(x) w_L(x) = 2^k - 2^{k-1}(1+1) = 0$ for $L = \emptyset$ and $-2^{k-1}\big((-1)+1\big) = 0$ for $|L| = 1$. For the lower bound, the empty set fails as in Proposition 5.1, and a singleton $\{y\}$ fails because $\varphi_y(x) = k - \sum_j y_j x_j$ lies in B_1^k, is non-negative, vanishes at y, and takes the value $2k > 0$ at the antipode of y. □

Proposition 5.2 is the control on the degree. The lower bound of Theorem 3.1 concerns $(B_{2m}^k)_+$, and at $q = 1$ there is no such m: indeed $u(k,1) = 2$, nowhere near $k+1$. Squaring is what raises the degree from 1 to 2, and without it the bound is false.

Proposition 5.3. *For every k and every $U \neq \mathcal{X}$, the set U is not a set of uniqueness for $(B_k^k)_+$. Hence $u(k,k) = 2^k$.*

Proof. The Walsh functions with $|L| \leq k$ are all of them, so by Lemma 2.1(2)(3) they form a basis of $\mathbb{R}^{\mathcal{X}}$ and $B_k^k = \mathbb{R}^{\mathcal{X}}$; explicitly $\delta_y = 2^{-k} \sum_L w_L(y)\, w_L$. If $y \notin U$ then δ_y is a nonzero non-negative member of B_k^k vanishing on U. The whole cube is trivially a set of uniqueness, so $u(k,k) = |\mathcal{X}| = 2^k$. □

This is Theorem 3.6(6) of [1]. It is the control against a too-large claim: no bound of the form $u(k,q) \leq \mathrm{poly}(k)$ can hold uniformly in q, so the fixed-q reading of Corollary 3.4 is essential.

Proposition 5.4. *For $k \geq 1$ the level set $W_{0,1} = \{\bot, u_1, \ldots, u_k\}$ has $k+1$ elements, and for $k \geq 2$ it is* not *a set of uniqueness for $(B_2^k)_+$.*

Proof. The subcube indicator $\varphi(x) = \frac{1}{4}(1+x_1)(1+x_2)$ expands as $\frac{1}{4}\big(w_\emptyset + w_{\{1\}} + w_{\{2\}} + w_{\{1,2\}}\big) \in B_2^k$ and is non-negative. Every point of $W_{0,1}$ has at most one coordinate equal to $+1$, so at least one of its first two coordinates is -1 and φ vanishes there; but $\varphi(\top) = 1$. □

Proposition 5.4 is the $D = \{0, 1\}$ instance of Theorem 2.5 of [1] — W_D is a set of uniqueness for the Ising class if and only if some $i, j \in D$ have $2 \le |i - j| \le k - 1$ — and the function exhibited is the one the source's own proof uses; we make no novelty claim for it. It is recorded because it shows that Theorem 4.2 is a statement about *which* $k + 1$ points: two level sets of the same cardinality $k + 1$ sit on opposite sides of the property.

6. Two corrections to [1]

Proposition 6.1. $u(2, 2) = 4 \neq 3$. *Hence the conjecture of* [1]*, in the form "$u(k, 2) = k + 1$ for every k", fails at $k = 2$, and the hypothesis $k \ge 3$ of Theorem 4.2 cannot be dropped.*

Proof. At $k = 2$ we have $q = 2 = k$, so $B_2^2 = \mathbb{R}^{\mathcal{X}}$ has dimension $1 + 2 + 1 = 4 = |\mathcal{X}|$ and Proposition 5.3 gives $u(2, 2) = 2^2 = 4$, whereas $k + 1 = 3$. □

The conjecture as printed has two clauses. Its second — no set of uniqueness has at most k elements — is Theorem 3.2 and is true for every k. Its first fails only at $k = 2$: the values $k = 0$ and $k = 1$ satisfy $u(k, 2) = k + 1$, since there $q = 2$ exceeds k and $u(k, 2) = u(k, k) = 2^k$, which is 1 and 2 respectively. So $k \ge 3$ is the correct hypothesis and $k = 2$ is the single exception.

The source's lower bounds rest on the covering number $g(k, q)$, and for $q = 2$ on a theorem of Kleitman and Spencer [3] quoted in its Lemma 3.3 as

$$g(k, 2) \;=\; \min\Big\{\, r \;:\; \binom{r-1}{\lfloor k/2 \rfloor - 1} \ge k \,\Big\}.$$

The floor should be $\lfloor r/2 \rfloor$.

Proposition 6.2. *The formula as printed is defective:*

1. *at $k = 3$ the set is empty — $\binom{r-1}{\lfloor 3/2 \rfloor - 1} = \binom{r-1}{0} = 1 < 3$ for every r — so the minimum does not exist;*
2. *at $k = 6$ it returns 5, since $\min\{r : \binom{r-1}{2} \ge 6\} = 5$.*

With $\lfloor r/2 \rfloor$ in place of $\lfloor k/2 \rfloor$ the formula returns $\min\{r : \binom{r-1}{\lfloor r/2 \rfloor - 1} \ge k\} = 4, 5, 6, 6$ for $k = 3, 4, 5, 6$.

Proof. All of these are finite arithmetic. For (1), $\lfloor 3/2 \rfloor - 1 = 0$ and $\binom{r-1}{0} = 1$ for every r. For (2) and the corrected values, evaluate: the sequence $\binom{r-1}{\lfloor r/2 \rfloor - 1}$ for $r = 2, 3, 4, 5, 6$ is $1, 1, 3, 4, 10$, and the least r with $\binom{r-1}{2} \ge 6$ is 5. □

Remark 6.3. The corrected values match the truth and the printed ones do not. Direct exhaustive search over the subsets of $\mathcal{X}$ meeting every $(k - 2)$-subcube gives

$$g(3, 2) = 4, \quad g(4, 2) = 5, \quad g(5, 2) = 6, \quad g(6, 2) = 6,$$

against the printed formula's undefined, $5, 6, 5$. Independently: writing the points of a candidate set as the rows of an $r \times k$ matrix of signs, the set meets every $(k-2)$-subcube exactly when its k columns are pairwise *qualitatively* 2*-independent*, that is, when every pair of columns realises all four sign patterns among the rows — a $(k-2)$-subcube being the set of points with two prescribed coordinates. Hence $g(k, 2) = \min\{r : M(r) \ge k\}$, where $M(r)$ is the largest number of pairwise qualitatively 2-independent columns of length r. An exact maximum-clique computation gives $M(r) = 1, 1, 3, 4, 10$ for $r = 2, \ldots, 6$, which is $\binom{r-1}{\lfloor r/2 \rfloor - 1}$ and not $\binom{r-1}{\lceil r/2 \rceil - 1}$, and inverting it returns $4, 5, 6, 6$ at $k = 3, 4, 5, 6$ as well. The $\lfloor r/2 \rfloor$ form also reproduces the asymptotic $\log k + \frac{1}{2} \log \log k + O(1)$ that [1] itself prints, while the printed form cannot: because its lower index varies with k and not with r, it grows linearly, returning 7 at $k = 10$, 12 at $k = 20$ and 52 at $k = 100$, where the $\lfloor r/2 \rfloor$ form returns 6, 8 and 10. Nothing in this remark belongs to the formal development: the values are computations and the identification above is elementary, and Proposition 6.2 claims only the arithmetic of the two formulae at the four values of k named there. Logarithms are to base 2: [1] declares no base, and no other one is possible, since $\binom{r-1}{\lfloor r/2 \rfloor - 1} \asymp 2^{r-1}/\sqrt{r}$, so that inverting the corrected formula gives

$r = \log_2 k + \frac{1}{2}\log_2\log_2 k + O(1)$, and a leading term $\log_b k = \log_2 k/\log_2 b$ matches that only for $b = 2$.

Remark 6.4. A third slip in [1], affecting nothing: the proof of its Theorem 3.4 describes $W_{\{1,k-1\}}$ as consisting of k elements, whereas $|W_{\{1,k-1\}}| = \binom{k}{1} + \binom{k}{k-1} = 2k$. The statement of that theorem, $u(k,3) \le 2k$, is the correct one, so nothing downstream is affected. We record it because $u(k,3)$ is the first quantity this paper leaves open: Theorem 3.3 gives $u(k,3) \ge k+1$ and [1] gives $u(k,3) \le 2k$.

7. Formal verification

Every numbered statement above — Lemmas 2.1 and 2.2, Theorems 3.1, 3.2, 3.3 and 4.2, Corollary 3.4's lower half, and Propositions 4.1, 5.1, 5.2, 5.3, 5.4, 6.1 and 6.2 — is formalized and machine-checked in Lean 4 [8] against Mathlib [9], with one caveat about Lemma 2.1(3): what the development proves is the linear independence of the Walsh system together with the count $\sum_{i\le q}\binom{k}{i}$ of the indices L with $|L| \le q$, which is what the proof of Theorem 3.1 uses, and not a separate statement identifying that count with the dimension of B_q^k as a subspace of $\mathbb{R}^{\mathcal{X}}$. The definitions of $\mathcal{X}$, of the Walsh functions, of B_q^k, of a set of uniqueness and of $u(k,q)$ are the development's own, transcribed from [1] as quoted in Section 1; $u(k,q)$ is defined as the infimum of the set of cardinalities of sets of uniqueness, and Theorem 4.2 is proved first in the form "$k+1$ is the least element of that set" and then in the form $u(k,2) = k+1$. The development is four files: the cube and the Walsh calculus with the certificate criterion; the lower bound; the certificate for $W_{1,k}$ and the main theorem; and the controls and corrections. There is no incomplete proof anywhere in it.

What rests on exhaustive computation: essentially nothing, which is worth stating precisely. Every statement listed above is proved symbolically and holds for all k (and, where relevant, all m and q); each of them has the axiom set {propositional extensionality, quotient soundness, choice}, and no compiled evaluator is invoked anywhere in the development. The sole exceptions are the two finite arithmetic facts inside Proposition 6.2, which are decided by kernel reduction over $r \le 6$; the corrected-formula half uses propositional extensionality alone. Nothing in the paper depends on a computation whose result the kernel does not check, with the single, labelled exception of the values of $g(k,2)$ in Remark 6.3, which are external computations and are stated there as such, and of the two upper bounds quoted from [1] (its Theorem 3.5, used in Corollary 3.4, and its Theorem 3.6(5), used in Remark 3.5), which are quoted and not formalized.

Independently of the formalization, the definitions were checked against the source's own printed values by an exact-rational decision procedure written from the definitions: U fails to be a set of uniqueness for $(B_q^k)_+$ exactly when the linear program $\{v \ge 0$ off $U,\ v = 0$ on $U,\ \langle v, w_L\rangle = 0$ for $|L| > q,\ \sum_x v(x) = 1\}$ is feasible, and feasibility was decided by a phase-one simplex over the rationals. With it we reproduced $u(k,0) = 1$ and $u(k,1) = 2$ for $k = 3,4$; $u(k,k) = 2^k$ for $k = 3,4$; the fact that $W_{1,k}$ is a set of uniqueness for $k = 3,4,5,6$; the fact that the level set of the even-weight points is one for $q = k-1$, $k = 3,4,5$; and Theorem 2.5 of [1] over *all* level sets W_D for $k = 3,4,5$ ($16 + 32 + 64$ sets, no mismatch). The claim of Theorem 3.2 was checked exhaustively for small k in two independent ways: for $k = 3,4,5$ every subset of size exactly k — 56, 1820 and 201 376 subsets respectively — was verified to admit the witness ℓ^2, with no failure; and for $k = 3,4$ the linear program was run on every one of those subsets, finding that none is a set of uniqueness. Proposition 6.1 was reproduced by running the linear program on all subsets of $\{-1,+1\}^2$ of size $0,1,2,3$. The explicit weights of Propositions 4.1 and 5.2 were evaluated in exact arithmetic for $3 \le k \le 9$ and $2 \le k \le 8$ respectively. Finally, at $k = 3$ exactly 2 of the 70 four-element subsets of the cube are sets of uniqueness — the two inscribed tetrahedra $W_{0,2}$ and $W_{1,3}$; classifying the extremal sets for general k is open.

On the novelty of Theorems 3.1, 3.2, 3.3 and 4.2. As of 3 September 2026 the source [1] had no citations in OpenAlex or Semantic Scholar, and no citing paper in a local full-text snapshot of arXiv sources; [2] was read in full and contains no bound on the size of a set of uniqueness other than the subcube one, and records $\dim B_q^k = \sum_{j\le q}\binom{k}{j}$ only as a fact, its Lemma 5.1, and never as a bound;

neither source contains the words *square*, *cubature*, *quadrature*, *orthogonal array*, *Rao* or *Delsarte*. A full-text sweep of that snapshot for *set(s) of uniqueness* returns 436 distinct papers, of which exactly two use the term in the sense of this paper, namely [1] and [2]; the rest use it in the classical Riemann–Cantor sense or in binary tomography. Web searches for the transplant of the cubature or Rao bound to sets of uniqueness on the cube returned nothing. What was *not* searched, and where a prior statement would most plausibly hide, is the pre-1990 design-theory book literature and the reviewing databases; we rate that risk low but not zero, and a reader who knows of such a statement should read Section 3 as a transplant rather than as a new argument, which is how it is described in Section 1 in any case.

The repository is private; repository reference to be added at submission.

References

[1] T. Skalski and T. Stroiński, *Level sets and maximum likelihood estimation for the Ising model*, arXiv:2511.20925v1 [math.ST], 25 November 2025, 13 pp. All quotations and all numbered references in this paper are to v1, read in full on 3 September 2026, on which date v1 was the only version and there was no journal reference. The theorem-like environments of the source share a single counter, so the numbering used here (Definition 1.1, Remark 1.3, Lemma 2.4, Theorem 2.5, Corollary 3.1, Lemma 3.3, Theorem 3.4, Theorem 3.5, Theorem 3.6, Conjecture 3.7) was read off a PDF compiled from the v1 e-print source rather than counted by hand.

[2] 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** (2022), 1045–1070; DOI 10.30757/ALEA.v19-43. Preprint arXiv: 1911.13143. Numbered references to it here, and in [1], are to the published version; the numbering of the current arXiv version differs (the dimension count is Lemma 5.1 in the published version and Lemma 3.18 in arXiv v3; the monotonicity remark is Remark 5.6 and Remark 3.22 respectively).

[3] D. J. Kleitman and J. Spencer, *Families of k-independent sets*, Discrete Math. **6** (1973), no. 3, 255–262; DOI 10.1016/0012-365X(73)90098-8. The formula displayed before Proposition 6.2 is attributed to this paper by [1], as its Theorem 1; we have not consulted it, and Proposition 6.2 and Remark 6.3 assert only arithmetic and finite computations of our own.

[4] C. R. Rao, *Factorial experiments derivable from combinatorial arrangements of arrays*, Supplement to the Journal of the Royal Statistical Society **9** (1947), no. 1, 128–139; DOI 10.2307/2983576. Cited for the mechanism of Section 3 — Rao's bound $N \geq \sum_{i \leq m} \binom{k}{i}$ for a binary orthogonal array of strength $2m$ on k factors — and not for any statement about sets of uniqueness.

[5] H. M. Möller, *Lower bounds for the number of nodes in cubature formulae*, in: Numerische Integration (G. Hämmerlin, ed.), Internat. Ser. Numer. Math. **45**, Birkhäuser, Basel, 1979, pp. 221–230; DOI 10.1007/978-3-0348-6288-2_17. The continuous form of the same mechanism: a positive cubature formula exact to degree $2m$ needs at least $\dim P_m$ nodes. In the cubature literature such lower bounds are conventionally credited to A. H. Stroud, *Quadrature methods for functions of more than one variable*, Ann. New York Acad. Sci. **86** (1960), 776–791; DOI 10.1111/j.1749-6632.1960.tb42842.x, which we have not consulted and on which nothing here depends.

[6] R. Cools, *Constructing cubature formulae: the science behind the art*, Acta Numerica **6** (1997), 1–54; DOI 10.1017/S0962492900002701.

[7] E. Stiemke, *Über positive Lösungen homogener linearer Gleichungen*, Math. Ann. **76** (1915), 340–342; DOI 10.1007/BF01458147. Cited only for the converse of Lemma 2.2, which is not used.

[8] L. de Moura and S. Ullrich, *The Lean 4 theorem prover and programming language*, in: Automated Deduction — CADE 28, Lecture Notes in Computer Science 12699, Springer, 2021, pp. 625–635; DOI 10.1007/978-3-030-79876-5_37.

[9] The mathlib Community, *The Lean mathematical library*, in: Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs (CPP 2020), ACM, 2020, pp. 367–381; DOI 10.1145/3372885.3373824.