

The Word-Search Constants $C_2(\text{ABBB}) = 8/5$ and $C_3(\text{ABB}) = 6$: On Two Questions of Schildkraut and of Halberstam and Schildkraut

Alper Ferudun*

October 10, 2026

Abstract

For a word w and a toroidal grid of letters in dimension d , an appearance of w is a cell together with one of the $3^d - 1$ directions in $\{-1, 0, 1\}^d \setminus \{0\}$ along which the letters spell w , and $C_d(w)$ is the supremum, over all grids, of the number of appearances per cell. We prove that $C_2(\text{ABBB}) = 8/5$. This answers Question 9.3 of Halberstam and Schildkraut (*Words with repeated letters in a grid*, 2025) and the last part of Question 11 of Schildkraut in the problem session of the Oberwolfach workshop *Combinatorics* (Oberwolfach Rep. 23 (2026)). We also prove that $C_3(\text{ABB}) = 6$, which is the case $d = 3$ of the first part of that Question 11 and of their Question 9.2. Both values also hold for $n \times \cdots \times n$ arrays without wraparound, as in the report. The lower bounds are known constructions. The upper bounds are computer-assisted: averaging lemmas, proved in the text, reduce them to 65,536 and 2^{27} integer inequalities, one for each pattern of a 4×4 or a $3 \times 3 \times 3$ window; these are given by exact certificates and were checked in integer arithmetic by independent programs. In both cases we classify the extremal grids. For ABBB they are the two five-queens lattices, on tori with both sides divisible by 5; for ABB in dimension 3 they are the grids in which the letters A form a family of parallel planes modulo 3. Every window that is not of this kind costs at least $3/70$, respectively $47137/294912$. So for these two pairs (d, w) , and only for these, we show that the supremum is attained by a periodic grid; Question 10 of the report (Question 2.5 of Halberstam and Schildkraut) asks this for all pairs. For ABBB we also determine the possible pairs (density of A, concentration): their closure is the quadrilateral with vertices $(0, 0)$, $(\frac{1}{5}, \frac{8}{5})$, $(\frac{3}{8}, \frac{3}{2})$, $(1, 0)$. It follows that Question 9.6 of Halberstam and Schildkraut, a stability statement with constant 1 for the letter distribution of near-extremal grids, has a negative answer for the pair $(\text{ABBB}, 2)$ as stated; it holds with the optimal constant $\frac{14}{5}$ for grids on two letters. Finally $18 \leq C_4(\text{ABB}) \leq \frac{62}{3}$, $54 \leq C_5(\text{ABB}) \leq 64$ and $\frac{24}{5} \leq C_3(\text{ABBB}) \leq \frac{11}{2}$. Question 10 in general and the values of $C_d(\text{ABB})$ for $d \geq 4$ remain open. This is an unrefereed note.

1 Introduction

1.1 The problem and the questions

Fill the cells of a large square array with letters and count how often a given word can be read as in a word-search puzzle: along rows, columns and diagonals, forwards and backwards. Patchell and Spiro [PS22] asked for the maximum of this count. Alon and Kravitz [AK24] determined its asymptotic behaviour in every dimension for words with distinct letters, and Halberstam and Schildkraut [HS25] studied words with repeated letters. We use the definitions of [HS25, §2.1].

A word $w = w_0 w_1 \cdots w_{\ell-1}$ is a finite string of letters with at least two distinct letters; throughout, A and B are two distinct letters. A grid of dimension d is a map $\Gamma: G \rightarrow \Sigma$, where $G = \mathbb{Z}/n_1 \times \cdots \times \mathbb{Z}/n_d$

*Mercury Software GmbH. Email: alper@mercurycodelab.com. GitHub: <https://github.com/AlperTheKing>.
Unrefereed preprint. CC BY 4.0.

for arbitrary positive integers $n_1, \dots, n_d$ and Σ is any set of letters; G is the *shape* and $|\Gamma| := |G|$ the *size* of Γ . Equivalently, Γ is a colouring of $\mathbb{Z}^d$ which is periodic under $n_1\mathbb{Z} \times \dots \times n_d\mathbb{Z}$. Let $D_d := \{-1, 0, 1\}^d \setminus \{0\}$ be the set of the $3^d - 1$ *directions*; for $d = 2$ these are the eight directions of a word search. An *appearance* of w in Γ is a pair $(p, v) \in G \times D_d$ such that $\Gamma(p + iv) = w_i$ for $i = 0, 1, \dots, \ell - 1$. The number of appearances is $\text{ct}(w, \Gamma)$, the *concentration* of w in Γ is $c_d(w, \Gamma) := \text{ct}(w, \Gamma)/|\Gamma|$ (it may exceed 1), and

$$C_d(w) := \sup\{c_d(w, \Gamma) : \Gamma \text{ a grid of dimension } d\}.$$

A grid with $c_d(w, \Gamma) = C_d(w)$ is called *w-extremal*. Three conventions should be kept in mind. An appearance is a start cell together with a direction; reversing a word, by $(p, v) \mapsto (p + (\ell - 1)v, -v)$, or renaming its letters does not change the constants, so ABBB , BBBA , BAAA , AAAB behave alike, and so do ABB , BBA , AAB , BAA . The sides n_i may be smaller than the length of the word, and may be 1; the cells $p + iv$ are taken modulo the shape and need not be distinct. All our bounds hold for every shape and every alphabet. Finally, the report [FKS26] counts the occurrences of w in an $n \times \dots \times n$ array without wraparound and writes their maximal number as $C_d(w)n^d - O(n^{d-1})$; for the two words of this note the two normalisations give the same constant (Corollary 1.2(b), Corollary 1.7).

The contribution of Schildkraut to the problem session of the workshop [FKS26, pp. 72–73] states the problem for arrays, recalls [PS22, AK24, HS25] and the value $C_2(\text{ABB}) = 2$ of [HS25], and asks two questions. Question 10 asks whether, for all d and w , the constant $C_d(w)$ is always “attained in the limit by some periodic construction” [FKS26, p. 72]. Question 11 asks for the value of $C_d(\text{ABB})$ for $d > 2$, and for the value of $C_2(\text{ABBB})$. The report adds that $C_2(\text{ABBB}) = 8/5$ is believed, and that $8/5$ is attained by placing A at the positions of five non-attacking queens in $(\mathbb{Z}/5)^2$ and B elsewhere.

These questions come from [HS25], where they are posed for grids. Question 2.5 there asks whether, for given w and d , a *w-extremal* grid exists, that is, whether the supremum defining $C_d(w)$ is attained. A colouring of $\mathbb{Z}^d$ which is periodic under a lattice of full rank is periodic under $m\mathbb{Z}^d$ for some m , hence is a grid, and a periodic colouring of concentration c has $cn^d - O(n^{d-1})$ occurrences in the n -array. In this sense Question 10 of the report is Question 2.5 of [HS25]. Question 9.2 asks whether ABB is *d-stackable* for every d , that is, whether $C_d(\text{ABB}) = 3^{d-1}C_1(\text{ABB}) = 2 \cdot 3^{d-2}$; the authors add that they do not know whether ABB is 3-stackable. Question 9.3 asks whether $C_2(\text{ABBB})$ equals $8/5$. In [HS25, §9] the pairs $(2, \text{ABBB})$ and $(3, \text{ABB})$ are named as two small cases in which Question 2.5 and the value of the constant are open.

Two further questions of [HS25] concern the *letter distribution* h_Γ of a grid, where $h_\Gamma(a)$ is the fraction of the cells with the letter a , and the total variation distance $\|h - h'\|_{\text{TV}} = \frac{1}{2} \sum_a |h(a) - h'(a)|$. Question 9.5 asks whether, for given w and d , all *w-extremal* grids of dimension d have the same letter distribution. Question 9.6 asks whether there is a letter distribution $h_{w,d}$ such that every grid Γ of dimension d with $c_d(w, \Gamma) \geq C_d(w)(1 - \delta)$ satisfies $\|h_\Gamma - h_{w,d}\|_{\text{TV}} \leq \delta$. They generalise two results of [HS25] for $d = 1$, Lemma 3.9 and Proposition 3.2 there; the latter is stated for every $\delta \geq 0$. The authors remark that this is a very strong quantitative form of such a statement, that a negative answer to Question 9.6 seems quite possible to them [HS25, §9], and that weaker versions would also be of interest.

The problem is record OWR-14299905-036 of the UnsolvedMath collection [ula26], where it is listed as open. The statement of the record consists of Questions 10 and 11 of the report, followed by a few lines which belong to the preceding problem of the report.

1.2 What was known

The following is due to [HS25] unless another source is given. For every word and every dimension,

$$3^{d-1}C_1(w) \leq C_d(w) \leq \frac{3^d - 1}{2} C_1(w) \quad [\text{HS25, Proposition 1.3}], \quad (1)$$

where the lower bound comes from *stacking*: $C_{d+1}(w) \geq 3C_d(w)$ [HS25, Lemma 2.2]. A word is d -stackable if the lower bound in (1) is an equality. There is a closed formula for $C_1(w)$ [HS25, Proposition 3.1]; in particular $C_1(ABB) = \frac{2}{3}$ and $C_1(ABBB) = \frac{1}{2}$ (see also Lemma 7.1). Alon and Kravitz [AK24] proved that $C_d(AB) = 3^{d-1}$ and that every word with distinct letters is d -stackable for all d ; their theorem is restated in this terminology in [HS25, Theorem 2.1]. A word of length at most four is 2-stackable unless it is of the form $ABBB$ or its reverse [HS25, Theorem 1.4(c)]. The crucial case is $C_2(ABB) = 2$ [HS25, Lemma 5.4], proved with a “local” linear-programming method [HS25, Proposition 5.2, §5.3], to which we return in Section 1.4.

For $ABBB$ in dimension 2 the best bounds were

$$\frac{8}{5} \leq C_2(ABBB) \leq \frac{59526}{35459} = 1.6787\dots;$$

the lower bound is the five-queens grid [HS25, Figure 5], and the upper bound was obtained with the local method and a computer check [HS25, Appendix A]. The same grid is extremal for two words of length five: $C_2(BABBB) = C_2(ABBBB) = \frac{8}{5}$ [HS25, Lemma 5.6, §6.2]. For ABB in dimension 3 the bounds (1) read $6 \leq C_3(ABB) \leq \frac{26}{3}$, and the bound $2 \cdot 3^{d-2}$ was known for grids of shape $(\mathbb{Z}/3)^d$ [HS25, Proposition 2.6], by Fourier analysis.

1.3 Results

Theorem 1.1. *Every grid Γ of dimension 2, of any shape and over any alphabet, satisfies*

$$\text{ct}(ABBB, \Gamma) \leq \frac{8}{5} |\Gamma|.$$

Equality holds for the grid of shape $(\mathbb{Z}/5)^2$ with A at the cells $(i, 2i \bmod 5)$ and B elsewhere. Hence $C_2(ABBB) = \frac{8}{5}$.

Corollary 1.2. (a) *If a word w contains $XXXX$ or $YYYY$ as a block of four consecutive letters, for two distinct letters X, Y , then $C_2(w) \leq \frac{8}{5}$.*

(b) *Let $M(n)$ be the maximal number of occurrences of $ABBB$ in an $n \times n$ array, in the eight directions and without wraparound. Then $M(n) \leq \frac{8}{5}n^2$ for all $n \geq 1$, and $M(n) \geq \frac{8}{5}(n-6)(n-10)$ for all $n \geq 6$. In particular $M(n) = \frac{8}{5}n^2 - O(n)$.*

Part (a) contains the upper bounds for $BABBB$ and $ABBBB$ of [HS25]; these do not imply Theorem 1.1.

For the next theorem let $x: \mathbb{Z}^2 \rightarrow \{0, 1\}$ be the periodic indicator function of the letter A in a grid Γ of shape G , and for $u \in \mathbb{Z}^2$ let X^u be the 4×4 window of x at u , that is, $X^u(r, c) := x(u + (r, c))$ for $0 \leq r, c \leq 3$. The *lattice windows* are the ten patterns

$$W_{s,t}(r, c) := [c \equiv sr + t \pmod{5}], \quad 0 \leq r, c \leq 3, \quad s \in \{2, 3\}, \quad t \in \mathbb{Z}/5, \quad (2)$$

where $[\cdot]$ is 1 if the statement in the brackets is true and 0 otherwise. Let $\mathcal{L}$ be the set of these ten windows.

Theorem 1.3 (extremal grids and a gap). (a) *A grid Γ of shape $\mathbb{Z}/n_1 \times \mathbb{Z}/n_2$ has the concentration $c_2(ABBB, \Gamma) = \frac{8}{5}$ if and only if both n_1 and n_2 are divisible by 5 and there are $s \in \{2, 3\}$ and $t \in \mathbb{Z}/5$ such that*

$$\Gamma(i, j) = \text{A} \quad \text{if } j \equiv si + t \pmod{5}, \quad \Gamma(i, j) = \text{B} \quad \text{otherwise.}$$

(b) For every grid Γ of dimension 2,

$$\frac{8}{5} - c_2(\text{ABBB}, \Gamma) \geq \frac{3}{70} \cdot \frac{\#\{u \in G : X^u \notin \mathcal{L}\}}{|G|}.$$

These are the five-queens grids of [HS25, Figure 5] and [FKS26, Figure 3]; the two values of s give mirror images of each other (Figure 2). So on a torus with both sides divisible by 5 there are exactly ten extremal grids, and on all other tori the value $\frac{8}{5}$ is not attained. In particular the supremum defining $C_2(\text{ABBB})$ is attained by a periodic grid, which is what Question 10 asks for the pair $(2, \text{ABBB})$.

Let $\alpha(\Gamma)$ be the density of the letter A in Γ , that is, $\alpha(\Gamma) = h_\Gamma(\text{A})$.

Theorem 1.4 (density of A). (a) Every grid Γ of dimension 2, with $\alpha = \alpha(\Gamma)$ and $c = c_2(\text{ABBB}, \Gamma)$, satisfies

$$c \leq 8\alpha, \quad c \leq \frac{12 - 4\alpha}{7}, \quad c \leq \frac{12}{5}(1 - \alpha).$$

(b) The closure of the set of all pairs $(\alpha(\Gamma), c_2(\text{ABBB}, \Gamma)) \in \mathbb{R}^2$, where Γ runs through the grids of dimension 2, is the quadrilateral Q with the vertices $(0, 0)$, $(\frac{1}{5}, \frac{8}{5})$, $(\frac{3}{8}, \frac{3}{2})$ and $(1, 0)$.

(c) If $c_2(\text{ABBB}, \Gamma) \geq \frac{8}{5}(1 - \delta)$ with $\delta \geq 0$, then

$$\frac{1 - \delta}{5} \leq \alpha(\Gamma) \leq \frac{1}{5} + \frac{14}{5}\delta,$$

and for no $\delta_0 > 0$ can the factor $\frac{14}{5}$ be replaced by a smaller number for all $\delta \in (0, \delta_0]$.

The vertices of Q are the pairs of the grid with the single letter B, of the five-queens grids, of the 4×4 grid with the rows AABA, BBAB, ABAB, ABAB, and of the grid with the single letter A (Figure 3).

Corollary 1.5 (Questions 9.5 and 9.6 of [HS25] for $(\text{ABBB}, 2)$). (a) All ABAB-extremal grids of dimension 2 have the letter distribution $h_* := (\text{A}: \frac{1}{5}, \text{B}: \frac{4}{5})$.

(b) There is no letter distribution h such that $\|h_\Gamma - h\|_{\text{TV}} \leq \delta$ holds for all $\delta \geq 0$ and all grids Γ of dimension 2 with $c_2(\text{ABBB}, \Gamma) \geq \frac{8}{5}(1 - \delta)$. More precisely, such an h would have to be h_* , and there are grids Γ_k on the letters A, B with $\delta_k := 1 - \frac{5}{8}c_2(\text{ABBB}, \Gamma_k) \rightarrow 0$ and $\|h_{\Gamma_k} - h_*\|_{\text{TV}} = \frac{112}{67}\delta_k$.

(c) On the other hand, if $c_2(\text{ABBB}, \Gamma) \geq \frac{8}{5}(1 - \delta)$, then $\|h_\Gamma - h_*\|_{\text{TV}} \leq \frac{14}{5}\delta$ for grids on the two letters A, B, and $\|h_\Gamma - h_*\|_{\text{TV}} \leq \frac{9002}{15}\delta$ for grids over any alphabet.

So, for the pair $(\text{ABBB}, 2)$, Question 9.5 has a positive answer, and Question 9.6 has a negative answer in the form in which it is stated, with the constant 1. What holds is a linear bound with a larger constant, a weaker version in the sense of [HS25]. By Theorem 1.4(c) the constant $\frac{14}{5}$ is optimal; the constant $\frac{9002}{15}$ is not optimised.

In dimension 3 let $x: \mathbb{Z}^3 \rightarrow \{0, 1\}$ again be the indicator of the letter A, let $X^u(y) := x(u + y)$ for $y \in \{0, 1, 2\}^3$ be the $3 \times 3 \times 3$ window of x at $u \in \mathbb{Z}^3$, and let $\mathcal{H}$ be the set of the 39 hyperplane windows

$$H_{a,c}(y) := [a \cdot y \equiv c \pmod{3}], \quad y \in \{0, 1, 2\}^3, \quad a \in D_3, \quad c \in \mathbb{Z}/3; \quad (3)$$

here $H_{a,c} = H_{-a,-c}$, and there are 13 vectors a up to sign.

Theorem 1.6. (a) Every grid Γ of dimension 3, of any shape and over any alphabet, satisfies $\text{ct}(\text{ABB}, \Gamma) \leq 6|\Gamma|$. Hence $C_3(\text{ABB}) = 6 = 3^2 C_1(\text{ABB})$, that is, ABB is 3-stackable.

- (b) A grid Γ of shape $\mathbb{Z}/n_1 \times \mathbb{Z}/n_2 \times \mathbb{Z}/n_3$ satisfies $c_3(\text{ABB}, \Gamma) = 6$ if and only if there are $a \in D_3$ and $c \in \mathbb{Z}/3$ such that 3 divides n_i whenever $a_i \neq 0$, and

$$\Gamma(y) = \text{A} \text{ if } a \cdot y \equiv c \pmod{3}, \quad \Gamma(y) = \text{B} \text{ otherwise.}$$

- (c) For every grid Γ of dimension 3 and shape G ,

$$6 - c_3(\text{ABB}, \Gamma) \geq \frac{47137}{294912} \cdot \frac{\#\{u \in G : X^u \notin \mathcal{H}\}}{|G|}.$$

The stacked grid of [HS25] is the case $a = (1, 0, 0)$. If exactly k of the sides are divisible by 3, there are $\frac{3}{2}(3^k - 1)$ extremal grids of that shape, so the supremum is attained by a periodic grid for the pair $(3, \text{ABB})$ as well. All extremal grids have the letter distribution $(\frac{1}{3}, \frac{2}{3})$; so Question 9.5 of [HS25] has a positive answer for the pair $(\text{ABB}, 3)$, too.

Corollary 1.7 (arrays in dimension 3). *Let $M_3(n)$ be the maximal number of occurrences of ABB in an $n \times n \times n$ array, in the 26 directions and without wraparound. Then $6(n-4)^2(n-6) \leq M_3(n) \leq 6n^3$ for all $n \geq 6$. In particular $M_3(n) = 6n^3 - O(n^2)$.*

Corollary 1.8 (higher dimensions, partial).

$$18 \leq C_4(\text{ABB}) \leq \frac{62}{3}, \quad 54 \leq C_5(\text{ABB}) \leq 64, \quad \frac{24}{5} \leq C_3(\text{ABBB}) \leq \frac{11}{2}.$$

The lower bounds in Corollary 1.8 are the stacking bounds; the upper bounds of (1) are $\frac{80}{3}$, $\frac{242}{3}$ and $\frac{13}{2}$. Table 1 collects the values.

word	d	C_d	source
AB	all d	3^{d-1}	[AK24]
ABB	1	$\frac{2}{3}$	[HS25, Proposition 3.1]
ABB	2	2	[HS25, Lemma 5.4]
ABB	3	6	Theorem 1.6; before: $6 \leq C_3 \leq \frac{26}{3}$ [HS25]
ABB	4	in $[18, \frac{62}{3}]$	Corollary 1.8; before: $[18, \frac{80}{3}]$ [HS25]
ABB	5	in $[54, 64]$	Corollary 1.8; before: $[54, \frac{242}{3}]$ [HS25]
ABBB	1	$\frac{1}{2}$	[HS25, Proposition 3.1]
ABBB	2	$\frac{8}{5}$	Theorem 1.1; before: $\frac{8}{5} \leq C_2 \leq \frac{59526}{35459}$ [HS25]
ABBB	3	in $[\frac{24}{5}, \frac{11}{2}]$	Corollary 1.8; before: $[\frac{24}{5}, \frac{13}{2}]$ [HS25]
BABBB, ABBBB	2	$\frac{8}{5}$	[HS25, Lemma 5.6, §6.2]

Table 1: The constants $C_d(w)$ for the words of this note. The two values in bold type are determined in this note; “before” refers to the bounds which follow from [HS25].

1.4 Method, and what stays open

The proofs of the two upper bounds have the same form. Only the positions of the letter **A** matter (Lemma 2.1). The number of appearances is the sum, over all positions of a window (4×4 , respectively $3 \times 3 \times 3$), of a weighted count Ψ_0 of the appearances inside the window (Lemma 2.2). To Ψ_0 one may add any function of the window whose sum over all positions vanishes for every grid, a *potential*: a difference of the values of one function on two overlapping blocks of the window (Lemma 2.4), or a

polynomial whose coefficients sum to zero on each class of congruent monomials (Lemma 2.5). If Ψ_0 plus a potential is at most M for every pattern of the window, then $C_d \leq M$. The last condition is a finite system of linear inequalities for the potential. We give potentials with rational values for which it holds with $M = \frac{8}{5}$ and $M = 6$; these are the *certificates*.

The lemmas just mentioned, the deduction of the theorems from the finite inequalities, the classification of the extremal grids from the list of the patterns with equality, and Corollaries 1.2, 1.5, 1.7 and 1.8 are proved in the text. The finite inequalities themselves, namely Propositions 3.1, 4.1, 5.1 and 6.1, are verified by computer, in exact integer arithmetic: 65,536 inequalities for each of the certificates in dimension 2 and $2^{27} = 134,217,728$ for each of the two certificates in dimension 3. Every certificate was checked by at least four different programs (Section 9); a check in dimension 2 takes about a second in pure Python, a check in dimension 3 between fifteen seconds and three minutes, depending on the program, and at most 2.5 gigabytes of memory. How the certificates were found plays no role in the proofs; it is described in Section 8.

The method strengthens the local method of [HS25, Proposition 5.2, §5.3]. There, a weighted count of the appearances in a finite set of cells is bounded pattern by pattern and then averaged over translations and symmetries; the unknowns of the linear program of [HS25, §5.3] are the weights, which are taken non-negative there. We use all potentials, that is, the full dual of the linear program over the distributions of windows with shift-invariant marginals (Section 8); this is what makes the 4×4 window sufficient for ABBB. Linear programming over locally consistent statistics, with dual certificates of this “discharging” type, is a standard technique: [HS25, §2.5] points to flag algebras [Raz07], an automated discharging method for density problems in grids is described in [Sto14], and bounds for ground-state energies of lattice models are obtained by convex optimisation in [HKD⁺16]. The technique is not new; as far as we know, its application to these two constants is (see “Scope and priority” at the end).

What stays open. Question 10 of the report, that is, Question 2.5 of [HS25], is settled here only for the pairs (2, ABBB) and (3, ABB). In general it remains open; the method gives attainment only when a finite window inequality happens to be tight and the patterns with equality fit together in a unique way. The first part of Question 11 and Question 9.2 of [HS25] are settled only for $d = 3$: the values $C_d(\text{ABB})$ for $d \geq 4$ remain open, with the bounds of Corollary 1.8 for $d = 4, 5$. Averaging over subspaces of dimension at most 3 cannot give $2 \cdot 3^{d-2}$ for $d \geq 4$ (Remark 7.3), and the window $\{0, 1, 2\}^4$ has 2^{81} patterns. Question 9.5 of [HS25] is answered only for these two pairs and Question 9.6 only for (ABBB, 2), and the best constant in Corollary 1.5(c) for alphabets with more than two letters is not determined.

2 Averaging and potentials

In this section $d \geq 1$ and $L \geq 1$ are integers, AB^L is the word consisting of one letter A followed by L letters B (so $L = 3$ for ABBB and $L = 2$ for ABB), and $\Gamma: G \rightarrow \Sigma$ is a grid of dimension d . Let $x = x_\Gamma: \mathbb{Z}^d \rightarrow \{0, 1\}$ be the periodic function with $x(p) = 1$ if and only if the letter of Γ at p , reduced modulo the shape, is A.

A *binary appearance* is a pair $(p, v) \in G \times D_d$ with $x(p) = 1$ and $x(p + kv) = 0$ for $k = 1, \dots, L$. Let $N_v = N_v(x)$ be the number of binary appearances with direction v , and $N = N(x) := \sum_{v \in D_d} N_v$.

Lemma 2.1. *Every appearance of AB^L in Γ is a binary appearance. Hence $\text{ct}(\text{AB}^L, \Gamma) \leq N(x)$, with equality if and only if every binary appearance is an appearance; this is the case if Γ has only the letters A and B.*

Proof. If (p, v) is an appearance, then $\Gamma(p) = \text{A}$ and $\Gamma(p + kv) = \text{B} \neq \text{A}$ for $1 \leq k \leq L$. □

So further letters can only destroy appearances, and it suffices to bound $N(x)$ for all periodic functions $x: \mathbb{Z}^d \rightarrow \{0, 1\}$.

Windows. Let $W := \{0, 1, \dots, L\}^d$. A *pattern* is an element $X \in \{0, 1\}^W$. An *arrow* of W is a pair $(c_0, v) \in W \times D_d$ with $c_0 + Lv \in W$; it *reads* AB^L in the pattern X if $X(c_0) = 1$ and $X(c_0 + kv) = 0$ for $1 \leq k \leq L$. For $v \in D_d$ let $|v|$ be the number of non-zero coordinates of v . The start cell of an arrow with direction v has the coordinate 0 where $v_i = 1$ and the coordinate L where $v_i = -1$, and its other coordinates are arbitrary. So there are $n_v = (L+1)^{d-|v|}$ arrows with direction v . We give each of them the weight $1/n_v$ and put

$$\Psi_0(X) := \sum_{\text{arrows } (c_0, v)} \frac{1}{n_v} [(c_0, v) \text{ reads } \text{AB}^L \text{ in } X]. \quad (4)$$

Thus, for every direction, the weights of the arrows with that direction sum to 1.

Let $\Pi = \Pi_d$ be the group of the $2^d d!$ signed permutations of the coordinates, acting linearly on $\mathbb{Z}^d$; it permutes D_d and preserves $|v|$. For $\sigma \in \Pi$ and $u \in \mathbb{Z}^d$ the *window* of x at u , turned by σ , is the pattern

$$X^{\sigma, u}(c) := x(u + \sigma(c)) \quad (c \in W), \quad X^u := X^{\text{id}, u}.$$

It depends only on u modulo the shape, so it is defined for $u \in G$.

Lemma 2.2 (averaging). *For every $\sigma \in \Pi$ we have $\sum_{u \in G} \Psi_0(X^{\sigma, u}) = N(x)$.*

Proof. Fix an arrow (c_0, v) . As σ is linear, $X^{\sigma, u}(c_0 + kv) = x(u + \sigma c_0 + k \sigma v)$. So the arrow reads AB^L in $X^{\sigma, u}$ if and only if $(u + \sigma c_0, \sigma v)$ is a binary appearance. When u runs through G , so does $u + \sigma c_0$; hence this happens for exactly $N_{\sigma v}$ elements $u \in G$. Summing over the arrows with their weights gives $\sum_u \Psi_0(X^{\sigma, u}) = \sum_v n_v \cdot \frac{1}{n_v} N_{\sigma v} = \sum_v N_{\sigma v} = N(x)$, because σ permutes D_d . $\square$

The windows are cut out of the periodic function x on $\mathbb{Z}^d$. If a side of the grid is shorter than $L+1$, several cells of a window are the same cell of G ; the proof does not use that they are different. So no condition on the shape is needed, here and below.

Potentials. A function $\Phi: \{0, 1\}^W \rightarrow \mathbb{Q}$ is a *potential* with respect to a non-empty subset $\Pi' \subseteq \Pi$ if

$$\sum_{\sigma \in \Pi'} \sum_{u \in G} \Phi(X^{\sigma, u}) = 0 \quad \text{for every grid of dimension } d.$$

Proposition 2.3. *Let Φ be a potential with respect to Π' , and let M be a number such that $\Psi_0(X) + \Phi(X) \leq M$ for all patterns $X \in \{0, 1\}^W$. Then every grid Γ of dimension d satisfies*

$$\text{ct}(\text{AB}^L, \Gamma) \leq N(x) = \frac{1}{|\Pi'|} \sum_{\sigma \in \Pi'} \sum_{u \in G} (\Psi_0 + \Phi)(X^{\sigma, u}) \leq M |G|. \quad (5)$$

In particular $C_d(\text{AB}^L) \leq M$. If $c_d(\text{AB}^L, \Gamma) = M$, then every binary appearance is an appearance, and $(\Psi_0 + \Phi)(X^{\sigma, u}) = M$ for all $\sigma \in \Pi'$ and $u \in G$.

Proof. The inequality on the left is Lemma 2.1, the equality is Lemma 2.2 together with the definition of a potential, and the inequality on the right holds term by term. If the two ends of (5) are equal, all inequalities are equalities. $\square$

We use two kinds of potentials. The first is for $d = 2$ and $L = 3$, so that $W = \{0, 1, 2, 3\}^2$; we write $X(r, c)$ with the row index r and the column index c . For a pattern X let

$$\text{top } X := (X(i, j))_{0 \leq i \leq 2, 0 \leq j \leq 3}, \quad \text{bot } X := (X(i + 1, j))_{0 \leq i \leq 2, 0 \leq j \leq 3}$$

be its upper and its lower 3×4 block, and let $X^t(r, c) := X(c, r)$ be the transposed pattern. For a function $g: \{0, 1\}^{3 \times 4} \rightarrow \mathbb{Q}$ put $D_g(X) := g(\text{top } X) - g(\text{bot } X)$.

Lemma 2.4 (telescoping). *For every $g: \{0, 1\}^{3 \times 4} \rightarrow \mathbb{Q}$ and every grid of dimension 2,*

$$\sum_{u \in G} D_g(X^u) = 0 \quad \text{and} \quad \sum_{u \in G} D_g((X^u)^t) = 0.$$

So $D_g + D_g^t$, where $D_g^t(X) := D_g(X^t)$, is a potential with respect to $\{\text{id}\}$.

Proof. We have $\text{bot } X^u = \text{top } X^{u+(1,0)}$, and $u \mapsto u + (1, 0)$ is a bijection of G ; this gives the first identity. For the second, $(X^u)^t(r, c) = x(u + (c, r))$, hence $\text{bot } (X^u)^t = \text{top } (X^{u+(0,1)})^t$, and $u \mapsto u + (0, 1)$ is a bijection of G . $\square$

The second kind is used for $d = 3$ and is stated for all d . Two finite sets $U, U' \subset \mathbb{Z}^d$ are *congruent* if $U' = \sigma(U) + t$ for some $\sigma \in \Pi$ and $t \in \mathbb{Z}^d$. Let $\lambda = (\lambda_U)$ be rational numbers indexed by the non-empty subsets U of W , and

$$Q_\lambda(X) := \sum_{\emptyset \neq U \subseteq W} \lambda_U \prod_{c \in U} X(c).$$

Lemma 2.5 (polynomial potentials). *Suppose that $\sum_{U \in \mathcal{C}} \lambda_U = 0$ for every congruence class $\mathcal{C}$ of non-empty subsets of W . Then Q_λ is a potential with respect to Π .*

Proof. For a finite set $V \subset \mathbb{Z}^d$ let $n(V) := \#\{u \in G : x(u + c) = 1 \text{ for all } c \in V\}$. Then $n(V + t) = n(V)$ for $t \in \mathbb{Z}^d$, and $\sum_{u \in G} \prod_{c \in U} X^{\sigma, u}(c) = n(\sigma U)$ for $U \subseteq W$. Put $S(U) := \sum_{\sigma \in \Pi} n(\sigma U)$. If $U' = \tau U + t$ with $\tau \in \Pi$, then $S(U') = \sum_{\sigma} n(\sigma \tau U + \sigma t) = \sum_{\sigma} n(\sigma \tau U) = S(U)$, since Π is a group. So S is constant on each congruence class, and

$$\sum_{\sigma \in \Pi} \sum_{u \in G} Q_\lambda(X^{\sigma, u}) = \sum_U \lambda_U S(U) = \sum_{\mathcal{C}} S(\mathcal{C}) \sum_{U \in \mathcal{C}} \lambda_U = 0. \quad \square$$

3 The constant $C_2(\text{ABBB})$

In this section and the next two, $d = 2$ and $L = 3$, so that $W = \{0, 1, 2, 3\}^2$. A *line* of a pattern $X \in \{0, 1\}^{4 \times 4}$ is one of its four rows, four columns and two diagonals. A line (z_0, z_1, z_2, z_3) carries two arrows, one from each end; at most one of them reads **ABBB**, and this happens if and only if the line is $(1, 0, 0, 0)$ or $(0, 0, 0, 1)$. We then say that the line reads **ABBB**. The window has 20 arrows (Figure 1), and (4) becomes

$$\Psi_0(X) = \frac{1}{4} \#\{\text{rows and columns of } X \text{ that read } \text{ABBB}\} + \#\{\text{diagonals of } X \text{ that read } \text{ABBB}\}.$$

Its largest value is $\frac{13}{4}$, so Ψ_0 alone proves nothing.

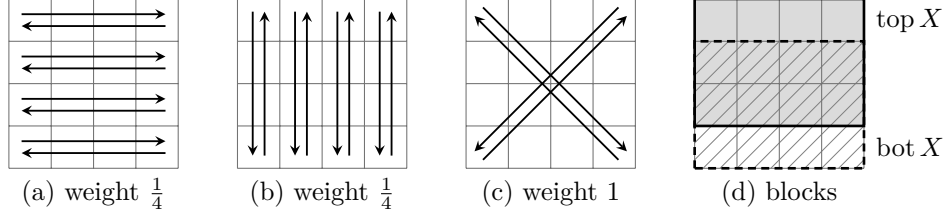


Figure 1: The 4×4 window. (a)–(c): its 20 arrows with their weights, each from a start cell to the third cell after it; for each of the eight directions the weights sum to 1, and Ψ_0 is the total weight of the arrows that read **ABBB**. (d): the blocks *top X* (rows 0, 1, 2; grey) and *bot X* (rows 1, 2, 3; hatched) of a pattern X .

The certificate. A function $g: \{0, 1\}^{3 \times 4} \rightarrow \mathbb{Q}$ is stored as a positive integer q and an integer table G with $g = G/q$. The 3×4 pattern y has the index

$$\text{ind}(y) := \sum_{(i,j): y(i,j)=1} 2^{4i+j} \in \{0, \dots, 4095\} \quad (0 \leq i \leq 2, 0 \leq j \leq 3).$$

A certificate file is a text file with comment lines beginning with **#**, one line with the letter **q** and the number q , and one line with the two numbers $\text{ind}(y)$ and $G(y)$ for each pattern y with $G(y) \neq 0$. The certificates of the second verification run are JSON files with the entries "**q**" and "**g**", the latter a list of the pairs $\text{ind}(y): G(y)$.

Proposition 3.1 (certificate for $\frac{8}{5}$; verified by computer). *Let $g = G/20$ be the function of the file `cert_ABBB_q20.txt` (1222 non-zero entries, $|G| \leq 29$). Then*

$$\Psi_0(X) + D_g(X) + D_g(X^t) \leq \frac{8}{5} \quad \text{for all } 65,536 \text{ patterns } X \in \{0, 1\}^{4 \times 4}. \quad (6)$$

The table is part of the reproducibility package and is not printed here. The program `check_gcert.py` (pure Python, standard library only) reads q and G and computes, for every pattern X , the integers $P(X) := 4\Psi_0(X)$, by counting the rows, columns and diagonals that read **ABBB**, and

$$\Delta(X) := G(\text{top } X) - G(\text{bot } X) + G(\text{top } X^t) - G(\text{bot } X^t),$$

directly from the definitions, and tests $5qP(X) + 20\Delta(X) \leq 32q$; this is (6) multiplied by $20q$. Only integers occur. For this certificate the maximum of the left side is $640 = 32q$, attained by 568 patterns. The same statement (6) holds for the two other certificates described in Proposition 4.1. The programs that checked (6) are listed in Section 9.

Proof of Theorem 1.1. By Lemma 2.4 the function $\Phi := D_g + D_g^t$ is a potential with respect to $\{\text{id}\}$, and $\Psi_0 + \Phi \leq \frac{8}{5}$ by Proposition 3.1. So Proposition 2.3 gives $\text{ct}(\text{ABBB}, \Gamma) \leq N(x) \leq \frac{8}{5}|\Gamma|$ for every grid of dimension 2.

For the lower bound let Γ_0 be the grid of shape $(\mathbb{Z}/5)^2$ with **A** at the cells (i, j) with $j \equiv 2i \pmod{5}$ and **B** elsewhere. Let p be an **A**-cell, $v \in D_2$ and $k \in \{1, 2, 3\}$. The cell $p + kv$ is an **A**-cell if and only if $k(v_2 - 2v_1) \equiv 0$, that is, $v_2 \equiv 2v_1 \pmod{5}$. This is impossible, because $v_2 - 2v_1 \in \{\pm 1, \pm 2, \pm 3\}$ for $v \in D_2$. So each of the five **A**-cells starts an appearance in each of the eight directions (Figure 2), and $c_2(\text{ABBB}, \Gamma_0) = \frac{40}{25} = \frac{8}{5}$. $\square$

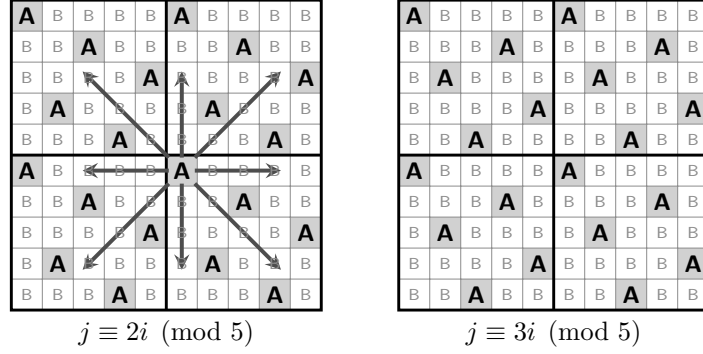


Figure 2: The two five-queens lattices (four copies of each extremal 5×5 grid; rows i numbered from the top, columns j from the left). Left: the eight appearances of $ABBB$ that start at one cell with the letter A . Right: the mirror image.

Proof of Corollary 1.2. (a) Let the block $XYYY$ occupy the positions $i, \dots, i+3$ of w . Then $(p, v) \mapsto (p+iv, v)$ maps the appearances of w in a grid Γ injectively to appearances of $XYYY$. If the block is $YYYX$, the map $(p, v) \mapsto (p+(i+3)v, -v)$ does the same. Hence $\text{ct}(w, \Gamma) \leq \text{ct}(XYYY, \Gamma) \leq \frac{8}{5}|\Gamma|$ by Theorem 1.1, applied with the letters X, Y in place of A, B . Here it matters that Theorem 1.1 holds for every alphabet, since w may have further letters.

(b) An occurrence in the $n \times n$ array is a pair (p, v) such that the cells $p, p+v, p+2v, p+3v$ lie in the array and carry the letters A, B, B, B . Read the array as a grid of shape $(\mathbb{Z}/n)^2$. Then every occurrence is an appearance, and different occurrences are different pairs. So $M(n) \leq \frac{8}{5}n^2$ by Theorem 1.1, for every $n \geq 1$.

For the lower bound let $n \geq 6$ and put the letter A at the cells (i, j) , $0 \leq i, j \leq n-1$, with $j \equiv 2i \pmod{5}$, and B elsewhere. If an A -cell satisfies $3 \leq i, j \leq n-4$, then the three cells after it in each of the eight directions lie in the array, and they carry the letter B by the computation in the proof of Theorem 1.1; so it starts eight occurrences. Each of the $n-6$ rows $i \in \{3, \dots, n-4\}$ contains at least $\lfloor (n-6)/5 \rfloor \geq (n-10)/5$ such cells, because among five consecutive columns exactly one satisfies the congruence. Hence $M(n) \geq 8(n-6)(n-10)/5$. (Some restriction on n is needed: an array with $n \leq 4$ has at most ten lines of four cells, each with at most one occurrence, so $M(n) \leq 10$, while $\frac{8}{5}(n-6)(n-10) \geq \frac{96}{5}$ for $n \leq 4$.) $\square$

4 Extremal grids and the gap

Proposition 4.1 (certificates with ten tight patterns; verified by computer). *Let $g' = G'/280$ be the function of the file `cert_ABBB_q280.txt` (1300 non-zero entries, $|G'| \leq 404$), and $\Psi := \Psi_0 + D_{g'} + D_{g'}^t$. Then*

$$\Psi(X) = \frac{8}{5} \quad \text{for the ten lattice windows } X \in \mathcal{L}, \quad \Psi(X) \leq \frac{8}{5} - \frac{3}{70} \quad \text{for all other patterns } X.$$

The same holds, with $\frac{53}{1260}$ in place of $\frac{3}{70}$, for the function $G''/2520$ of the file `mycert_ABBB.json` (3756 non-zero entries), which was produced by the second verification run with its own linear program.

So there are three certificates for (6), found by two different computations. We now collect the properties of the lattice windows (2) which are needed.

Lemma 4.2. (a) *For $s \in \{2, 3\}$ and $t \in \mathbb{Z}/5$ let $y_{s,t}(i, j) := [j \equiv si + t \pmod{5}]$ for $(i, j) \in \mathbb{Z}^2$. Every 4×4 window of $y_{s,t}$ is a lattice window; the window at (a, b) is $W_{s,t+sa-b}$. In particular $W_{s,t}(r, c+1) = W_{s,t-1}(r, c)$ and $W_{s,t}(r+1, c) = W_{s,t+s}(r, c)$.*

- (b) *The ten lattice windows are distinct. If $T' \in \mathcal{L}$ and $T'(r, c) = W_{s,t}(r, c+1)$ for all $0 \leq r \leq 3$ and $0 \leq c \leq 2$, then $T' = W_{s,t-1}$. If $T' \in \mathcal{L}$ and $T'(r, c) = W_{s,t}(r+1, c)$ for all $0 \leq r \leq 2$ and $0 \leq c \leq 3$, then $T' = W_{s,t+s}$.*
- (c) *In $y_{s,t}$, if $y_{s,t}(p) = 1$, then $y_{s,t}(p+kv) = 0$ for all $v \in D_2$ and $k \in \{1, 2, 3\}$. If $y_{s,t}(m) = 0$, then there are a cell p in the row of m and $v \in \{(0, 1), (0, -1)\}$ such that $y_{s,t}(p) = 1$ and m is one of the cells $p+v, p+2v, p+3v$.*

Proof. (a) $y_{s,t}(a+r, b+c) = [c \equiv sr + (t+sa-b)]$.

(b) The rows $r = 0, 1, 2, 3$ of $W_{s,t}$ have their entry 1, if any, in the column $c \in \{0, 1, 2, 3\}$ with $c \equiv t+sr \pmod{5}$. The four residues $t+sr$ are distinct, because s is invertible modulo 5; so at most one of them is 4, and every lattice window has at least three entries 1, in different rows. A lattice window determines (s, t) as soon as two of its entries 1 are known, say at (r, c) and (r', c') with $r \neq r'$: then $s(r-r') \equiv c-c'$, and $r-r' \in \{\pm 1, \pm 2, \pm 3\}$ is invertible modulo 5, so s is determined, and $t \equiv c-sr$. In particular the ten windows are distinct. For the second statement note that at most two of the four residues lie in $\{0, 4\}$, so at least two rows of $W_{s,t}$ have their entry 1 in the columns 1, 2, 3. These two entries are entries of T' , one column to the left. By what was said, T' is determined by them, and $W_{s,t-1}$ has them by (a). For the third statement, at most one of the three residues $t+sr$, $r \in \{1, 2, 3\}$, is equal to 4, so at least two of the rows 1, 2, 3 of $W_{s,t}$ contain an entry 1, and the same argument applies with $W_{s,t+s}$.

(c) $y_{s,t}(p+kv) = 1$ would mean $k(v_2 - sv_1) \equiv 0$, hence $v_2 \equiv sv_1 \pmod{5}$; but for $v \in D_2$ and $s \in \{2, 3\}$ the number $v_2 - sv_1$ is not divisible by 5 (for $s = 3$ it lies in $\{\pm 1, \pm 2, \pm 3, \pm 4\}$). Let $y_{s,t}(m) = 0$, $m = (i, j)$, and let $\rho \in \{1, 2, 3, 4\}$ be the residue of $j - si - t$. If $\rho \leq 3$ take $p = (i, j - \rho)$ and $v = (0, 1)$; if $\rho = 4$ take $p = (i, j + 1)$ and $v = (0, -1)$. $\square$

Proof of Theorem 1.3. (b) By Proposition 2.3 with the potential $D_{g'} + D_{g'}^t$, and by Proposition 4.1,

$$\text{ct}(\text{ABBB}, \Gamma) \leq N(x) = \sum_{u \in G} \Psi(X^u) \leq \frac{8}{5}|G| - \frac{3}{70} \#\{u \in G : X^u \notin \mathcal{L}\}.$$

(a) Let $c_2(\text{ABBB}, \Gamma) = \frac{8}{5}$. By (b), or by Proposition 2.3, every binary appearance is an appearance and $X^u \in \mathcal{L}$ for all $u \in G$, hence for all $u \in \mathbb{Z}^2$. Let $X^{(0,0)} = W_{s,t_0}$. Since $X^{u+(0,1)}(r, c) = X^u(r, c+1)$ and $X^{u+(1,0)}(r, c) = X^u(r+1, c)$, Lemma 4.2(b) and induction give $X^{(a,b)} = W_{s,t_0+sa-b}$ for all integers $a, b \geq 0$. Therefore

$$x(a, b) = X^{(a,b)}(0, 0) = [b \equiv sa + t_0 \pmod{5}] \quad (a, b \geq 0).$$

The function x is periodic under $(n_1, 0)$ and $(0, n_2)$. Applied at points with $a, b \geq 0$ this forces $sn_1 \equiv 0$ and $n_2 \equiv 0 \pmod{5}$, that is, $5 \mid n_1$ and $5 \mid n_2$. Then y_{s,t_0} has these periods as well and agrees with x on a set of representatives, so $x = y_{s,t_0}$ on $\mathbb{Z}^2$. Finally let m be a cell with $x(m) = 0$. By Lemma 4.2(c) it is one of the cells $p+v, p+2v, p+3v$ of a binary appearance (p, v) . This is an appearance, so the letter at m is B. Hence Γ has the form stated in the theorem.

Conversely, let $5 \mid n_1$, $5 \mid n_2$, and let Γ have the letter A at the cells with $j \equiv si + t$ and B elsewhere. The density of A is $\frac{1}{5}$, and by Lemma 4.2(c) every A-cell starts eight appearances. So $c_2(\text{ABBB}, \Gamma) = \frac{8}{5}$. $\square$

5 The density of the letter A and Question 9.6

For a pattern $X \in \{0, 1\}^{4 \times 4}$ let $n_A(X)$ be the number of its entries equal to 1.

Proposition 5.1 (certificates for the density bounds; verified by computer). (a) Let $g = G/280$ be the function of the file `cert_ABBB_density_q280.txt` (1628 non-zero entries, $|G| \leq 431$). Then, for all 65,536 patterns X ,

$$\Psi_0(X) + \frac{4}{7} \cdot \frac{n_A(X)}{16} + D_g(X) + D_g(X^t) \leq \frac{12}{7}. \quad (7)$$

The same holds for the function of the file `mycert_ABBB_density.json` of the second verification run.

(b) Let $g = G/2520$ be the function of the file `mycert_ABBB_density_piece3.json` of the second verification run (3792 non-zero entries). Then, for all 65,536 patterns X ,

$$\Psi_0(X) + \frac{12}{5} \cdot \frac{n_A(X)}{16} + D_g(X) + D_g(X^t) \leq \frac{12}{5}. \quad (8)$$

Equality holds in (7) for 206 patterns, for both certificates, and in (8) for 65 patterns, among them the pattern with all entries 1.

Proof of Theorem 1.4(a). Every appearance starts at a cell with the letter A, and a cell starts at most eight appearances; so $c \leq 8\alpha$. For the other two bounds note that $\sum_{u \in G} n_A(X^u) = \sum_{c \in W} \sum_{u \in G} x(u+c) = 16\alpha|G|$. Let (κ, R) be $(\frac{4}{7}, \frac{12}{7})$ or $(\frac{12}{5}, \frac{12}{5})$, and let g be the corresponding function of Proposition 5.1. By Lemmas 2.1, 2.2 and 2.4,

$$\text{ct}(\text{ABBB}, \Gamma) + \kappa\alpha|G| \leq \sum_{u \in G} \left(\Psi_0 + \kappa \frac{n_A}{16} + D_g + D_g^t \right)(X^u) \leq R|G|.$$

Hence $c \leq R - \kappa\alpha$, which is $c \leq \frac{12-4\alpha}{7}$ and $c \leq \frac{12}{5}(1-\alpha)$. $\square$

Together with $c \geq 0$ the three bounds describe the quadrilateral Q (Figure 3). The first two alone give $c \leq \frac{8}{5}$, so (7) gives another proof of the upper bound in Theorem 1.1.

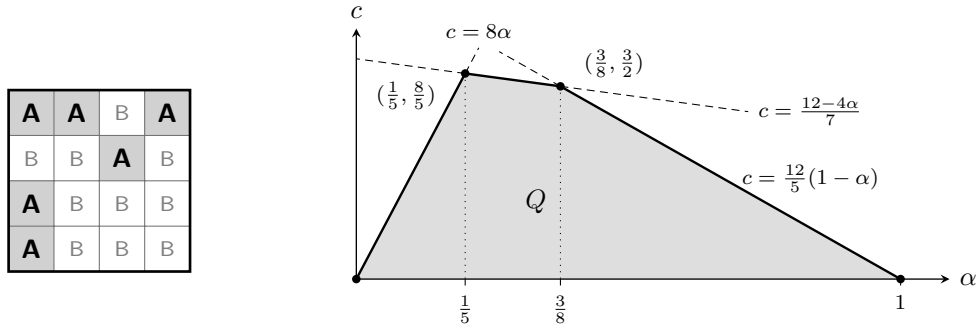


Figure 3: Left: the 4×4 grid with the rows AABA, BBAB, ABBB, ABBB; it has 24 appearances of ABBB, density $\frac{3}{8}$ of A and concentration $\frac{3}{2}$. Right: the quadrilateral Q of Theorem 1.4 in the plane of the pairs (density α of A, concentration c), with the three lines of Theorem 1.4(a).

Lemma 5.2 (mixing). Let Γ_1, Γ_2 be grids of dimension 2 with the shapes $\mathbb{Z}/m_1 \times \mathbb{Z}/n_1$ and $\mathbb{Z}/m_2 \times \mathbb{Z}/n_2$, and with the pairs $(\alpha_i, c^{(i)}) := (\alpha(\Gamma_i), c_2(\text{ABBB}, \Gamma_i))$. Let m be the least common multiple of m_1, m_2 , let L be a positive multiple of n_1 and $N > L$ such that $N - L$ is a multiple of n_2 , and let Γ be the grid of shape $\mathbb{Z}/m \times \mathbb{Z}/N$ with

$$\Gamma(i, j) := \begin{cases} \Gamma_1(i \bmod m_1, j \bmod n_1) & \text{for } 0 \leq j < L, \\ \Gamma_2(i \bmod m_2, (j - L) \bmod n_2) & \text{for } L \leq j < N. \end{cases}$$

Then, with $\lambda := L/N$,

$$\alpha(\Gamma) = \lambda\alpha_1 + (1 - \lambda)\alpha_2, \quad |c_2(\text{ABBB}, \Gamma) - \lambda c^{(1)} - (1 - \lambda)c^{(2)}| \leq \frac{96}{N}.$$

Consequently the closure of the set of all pairs $(\alpha(\Gamma), c_2(\text{ABBB}, \Gamma))$ is convex.

Proof. The statement on α is clear. Let a_j be the number of appearances (p, v) in Γ with p in the column j . Whether (p, v) is an appearance depends only on the letters in the columns $j - 3, \dots, j + 3$. If these columns all lie in $\{0, \dots, L - 1\}$, then Γ agrees there with the periodic extension Γ'_1 of Γ_1 to the shape $\mathbb{Z}/m \times \mathbb{Z}/L$, and a_j is the number a'_j of appearances in Γ'_1 which start in the column j . The grid Γ'_1 has the concentration $c^{(1)}$, because m_1 divides m and n_1 divides L ; so the sum of a'_j over all columns $0 \leq j < L$ is $mLc^{(1)}$. Moreover $0 \leq a_j, a'_j \leq 8m$. The same holds for the columns $L \leq j < N$ and Γ_2 . At most six columns $j < L$ and at most six columns $j \geq L$ are not of the kind just described. Hence $|\text{ct}(\text{ABBB}, \Gamma) - mLc^{(1)} - m(N - L)c^{(2)}| \leq 12 \cdot 8m$, and division by mN gives the estimate.

For the last statement note that replacing (L, N) by (kL, kN) does not change λ and makes $96/N$ arbitrarily small, and that the admissible numbers L/N are dense in $[0, 1]$. So the closure contains the segment between the pairs of any two grids, and hence is convex. $\square$

Proof of Theorem 1.4(b). By part (a), and since $c \geq 0$ and $0 \leq \alpha \leq 1$, all pairs lie in Q . The four vertices of Q are the pairs of the grids named after the theorem: this is clear for the two grids with a single letter, it is Theorem 1.1 for the five-queens grid, and the 4×4 grid of Figure 3 has six letters A and 24 appearances (counted by the programs of Section 9). By Lemma 5.2 the closure contains the convex hull of these four points, which is Q . $\square$

The next family, found by the first verification run, lies near the side of Q between $(\frac{1}{5}, \frac{8}{5})$ and $(\frac{3}{8}, \frac{3}{2})$.

Proposition 5.3. *Let P be the 4×4 grid with the rows AAAB, BBAB, BBBA, BBAB; it is a translate of the transpose of the grid of Figure 3. For multiples L, N of 20 with $20 \leq L \leq N - 20$ let $\Gamma_{L,N}$ be the grid of shape $\mathbb{Z}/20 \times \mathbb{Z}/N$ with*

$$\Gamma_{L,N}(i, j) := \begin{cases} P(i \bmod 4, j \bmod 4) & \text{for } 0 \leq j < L, \\ \text{A} & \text{for } L \leq j < N \text{ and } j \equiv 2i \pmod{5}, \\ \text{B} & \text{for } L \leq j < N \text{ and } j \not\equiv 2i \pmod{5}. \end{cases}$$

Then $\Gamma_{L,N}$ has $4N + \frac{7}{2}L$ letters A and

$$\text{ct}(\text{ABBB}, \Gamma_{L,N}) = 32N - 2L - 54.$$

Hence, with $\delta := 1 - \frac{5}{8}c_2(\text{ABBB}, \Gamma_{L,N})$ and $\alpha := \alpha(\Gamma_{L,N})$,

$$\delta = \frac{2L + 54}{32N}, \quad \alpha - \frac{1}{5} = \frac{7L}{40N}, \quad \frac{\alpha - \frac{1}{5}}{\delta} = \frac{28L}{5(2L + 54)} > 1. \quad (9)$$

Proof. The strip of the columns $j < L$ has $20L$ cells, of which $\frac{3}{8}$ carry the letter A, and the other part has $20(N - L)$ cells, of which $\frac{1}{5}$ carry A. For the count we argue as in the proof of Lemma 5.2. Let a_j be the number of appearances starting in the column j . For $3 \leq j \leq L - 4$ it equals the corresponding number $s_{j \bmod 4}$ for the 4-periodic colouring given by P , and $s_0 + s_1 + s_2 + s_3 = 5 \cdot 24$. For $L + 3 \leq j \leq N - 4$ it equals 32, because a column of the lattice has four cells with the letter A, each of which starts eight appearances. For the twelve other columns, $j = L + k$ and $j \equiv N + k \pmod{N}$

with $-3 \leq k \leq 2$, the number a_j depends only on the letters in the columns $L + k'$, respectively $N + k'$, with $-6 \leq k' \leq 5$. Since L and N are multiples of 20, and $L \geq 6$, $N - L \geq 6$, these letters do not depend on L and N . Summing up,

$$\text{ct}(\text{ABBB}, \Gamma_{L,N}) = 30L - (s_0 + 2s_1 + 2s_2 + s_3) + 32(N - L - 6) + \kappa_1 = 32N - 2L - \kappa$$

with constants κ_1, κ that do not depend on L and N . The grid $\Gamma_{20,40}$ has 1186 appearances, counted from the definition by the program `strip_family.py`; so $\kappa = 54$. (The formula also agrees with the counts of 19 further grids of the family by the same program, with the 20 counts of the first verification run, for which $L \leq 400$ and $N \leq 100,000$, and with the counts of all 190 grids of the family with $N \leq 400$ by the third verification run.) The formulas (9) follow, and $28L > 5(2L + 54)$ for $L \geq 20$. $\square$

Proof of Theorem 1.4(c). By part (a), $\alpha \geq \frac{c}{8} \geq \frac{1-\delta}{5}$ and $\alpha \leq 3 - \frac{7}{4}c \leq 3 - \frac{14}{5}(1 - \delta) = \frac{1}{5} + \frac{14}{5}\delta$. For the grids $\Gamma_{L,N}$ of Proposition 5.3 the quotient $(\alpha - \frac{1}{5})/\delta$ tends to $\frac{14}{5}$ as $L \rightarrow \infty$, while δ is as small as we like if N/L is large. $\square$

Proof of Corollary 1.5. (a) This is part of Theorem 1.3(a).

(b) Let h have the property in question. The five-queens grid Γ_0 has $c_2(\text{ABBB}, \Gamma_0) = \frac{8}{5} = C_2(\text{ABBB})$ by Theorem 1.1. So the property with $\delta = 0$ gives $h = h_{\Gamma_0} = h_*$. Let $\Gamma_k := \Gamma_{40,20k}$ for $k \geq 3$. By (9), $\delta_k = \frac{67}{320k} \rightarrow 0$ and $\|h_{\Gamma_k} - h_*\|_{\text{TV}} = \alpha(\Gamma_k) - \frac{1}{5} = \frac{7}{20k} = \frac{112}{67}\delta_k > \delta_k$. This contradicts the property with $\delta = \delta_k$. (The single grid of Figure 3 already gives a contradiction: there $\delta = \frac{1}{16}$ and the distance is $\frac{7}{40}$.)

(c) We may assume $\delta = 1 - \frac{5}{8}c_2(\text{ABBB}, \Gamma)$. For a grid on the letters **A, B** the distance is $|\alpha - \frac{1}{5}|$, which is at most $\frac{14}{5}\delta$ by Theorem 1.4(c). In general let β and γ be the densities of the letter **B** and of the cells with a letter different from **A** and **B**; we call the latter cells *foreign*. We show $\gamma \leq \frac{1792}{3}\delta$. Since $\beta - \frac{4}{5} = \frac{1}{5} - \alpha - \gamma$, this gives

$$\|h_\Gamma - h_*\|_{\text{TV}} = \frac{1}{2}(|\alpha - \frac{1}{5}| + |\beta - \frac{4}{5}| + \gamma) \leq |\alpha - \frac{1}{5}| + \gamma \leq (\frac{14}{5} + \frac{1792}{3})\delta = \frac{9002}{15}\delta.$$

Let b be the number of $u \in G$ with $X^u \notin \mathcal{L}$, and $e := N(x) - \text{ct}(\text{ABBB}, \Gamma)$ the number of binary appearances which are not appearances. Let m be a foreign cell, so $x(m) = 0$. Suppose that the sixteen windows $X^{m-(r,c)}$, $0 \leq r, c \leq 3$, which contain m , are lattice windows. By Lemma 4.2(a),(b) they are windows of one configuration $y_{s,t}$, so x agrees with a translate of $y_{s,t}$ on the 7×7 square with the centre m . By Lemma 4.2(c) there is a binary appearance (p, v) of x , in the row of m and inside this square, with $m \in \{p+v, p+2v, p+3v\}$. It is not an appearance, because the letter at m is not **B**. So every foreign cell lies in a window $X^u \notin \mathcal{L}$, or on a binary appearance which is not an appearance. There are at most $16b$ cells of the first kind and at most $3e$ of the second. By the inequality in the proof of Theorem 1.3(b), $b \leq \frac{70}{3}(\frac{8}{5}|G| - N(x))$. Hence the number of foreign cells is at most

$$\frac{1120}{3}(\frac{8}{5}|G| - N(x)) + 3(N(x) - \text{ct}(\text{ABBB}, \Gamma)) \leq \frac{1120}{3}(\frac{8}{5}|G| - \text{ct}(\text{ABBB}, \Gamma)) = \frac{1792}{3}\delta |G|. \quad \square$$

6 Dimension three: $C_3(\text{ABB}) = 6$

Now $d = 3$ and $L = 2$, so that $W = \{0, 1, 2\}^3$ has 27 cells and 2^{27} patterns, and Π is the group of the 48 signed permutations of the coordinates. A *line* of W is a set $\{m - v, m, m + v\} \subset W$ with $v \in D_3$; it exists if and only if $m_i = 1$ whenever $v_i \neq 0$. There are 27 lines parallel to an axis, 18 parallel to a diagonal of a face and 4 space diagonals, 49 in all. A line carries the two arrows $(m - v, v)$ and

$(m + v, -v)$, and exactly one of them reads ABB if $X(m) = 0$ and $X(m - v) \neq X(m + v)$; otherwise none does. So (4) becomes

$$\Psi_0(X) = \sum_{\text{lines } \{m-v, m, m+v\}} w_v [X(m) = 0] [X(m - v) \neq X(m + v)], \quad (10)$$

where $w_v = \frac{1}{9}, \frac{1}{3}, 1$ for $|v| = 1, 2, 3$.

The certificates. For $\sigma \in \Pi$ let $\sigma_W(y) := \sigma(y - m_0) + m_0$ with $m_0 := (1, 1, 1)$; these are the 48 symmetries of the cube W . A certificate is a JSON file with a positive integer "Qden", divisible by 9, and a list "terms"; each term consists of a "support" S , a set of one to four cells of W , and an integer "Z". It defines the coefficients

$$\lambda_U := \frac{Z_S}{\text{Qden}} \quad \text{if } U = \sigma_W(S) \text{ for a listed support } S \text{ and some } \sigma \in \Pi, \quad \lambda_U := 0 \quad \text{otherwise,}$$

and so a polynomial Q_λ of degree at most 4 in the 27 variables $X(c)$, which is invariant under the symmetries of the cube.

Proposition 6.1 (certificates for 6; verified by computer). *Let λ be given by the file `cert_ABB_3D.json` (160 supports, of which 4, 10, 26, 120 have 1, 2, 3, 4 cells; Qden = 1,451,520; 4941 monomials), or by the file `mycert_ABB_3D.json` of the second verification run (437 supports, of which 4, 18, 88, 327 have 1, 2, 3, 4 cells; Qden = 589,824; 15,245 monomials). In both cases:*

- (a) *no monomial arises from two listed supports, and $\sum_{U \in \mathcal{C}} \lambda_U = 0$ for every congruence class $\mathcal{C}$ of non-empty subsets of W ;*
- (b) *$\Psi_0(X) + Q_\lambda(X) \leq 6$ for all $2^{27} = 134,217,728$ patterns $X \in \{0, 1\}^W$;*
- (c) *equality holds in (b) exactly for the 39 hyperplane windows $X \in \mathcal{H}$ of (3), and all other patterns satisfy $\Psi_0(X) + Q_\lambda(X) \leq 6 - \eta$, with $\eta = \frac{1079}{90720}$ for the first file and $\eta = \frac{47137}{294912}$ for the second.*

The certificate meets 73, respectively 177, of the 348 congruence classes of subsets of W with at most four cells. The program `check_abb3d.py` does the following, with integers only. It expands the list of terms into monomials and stops if a monomial arises twice. It computes for every listed support the canonical representative of its congruence class, namely the lexicographically smallest of its 48 images under Π , each translated so that the least value of every coordinate is 0, and verifies that the coefficients in each class sum to zero. It verifies a bound which excludes an overflow of 64-bit integers. It computes $F(X) := \sum_{U \subseteq X} Z_U$ for all 2^{27} patterns at once by the subset-sum transform (27 passes over a table of 2^{27} integers), computes $9\Psi_0(X)$ from the 49 lines of (10), and tests

$$9F(X) + \text{Qden} \cdot 9\Psi_0(X) \leq 54 \cdot \text{Qden},$$

which is (b) multiplied by $9 \cdot \text{Qden}$. It records the patterns with equality and the largest value below $54 \cdot \text{Qden}$, compares the former with the 39 windows $H_{a,c}$, tests the statement of Lemma 6.2(b) below, and counts the appearances in the grid Γ_1 below from the definition. The other programs that checked Proposition 6.1 are described in Section 9.

Proof of Theorem 1.6(a). By Proposition 6.1(a) and Lemma 2.5, Q_λ is a potential with respect to Π , and $\Psi_0 + Q_\lambda \leq 6$ by Proposition 6.1(b). So Proposition 2.3 gives $\text{ct}(\text{ABB}, \Gamma) \leq N(x) \leq 6|\Gamma|$ for every grid of dimension 3. Let Γ_1 be the grid of shape $(\mathbb{Z}/3)^3$ with A at the nine cells with $y_1 = 0$ and B elsewhere. Each A-cell starts an appearance in each of the 18 directions v with $v_1 \neq 0$, so Γ_1 has $162 = 6 \cdot 27$ appearances. Since $C_1(\text{ABB}) = \frac{2}{3}$, this is the statement that ABB is 3-stackable. $\square$

Hyperplane windows. The window $W = \{0, 1, 2\}^3$ is a set of representatives of $(\mathbb{Z}/3)^3 = \mathbb{F}_3^3$, and under this identification the hyperplane windows $H_{a,c}$ of (3) are the indicator functions of the 39 affine planes of $\mathbb{F}_3^3$.

Lemma 6.2. (a) For $a \in D_3$ and $c \in \mathbb{Z}/3$ let $y_{a,c}(p) := [a \cdot p \equiv c \pmod{3}]$ for $p \in \mathbb{Z}^3$. The window of $y_{a,c}$ at u is $H_{a,c-a \cdot u}$. In particular $H_{a,c}(y + e_i) = H_{a,c-a_i}(y)$ for the unit vectors e_i .

(b) Let $i \in \{1, 2, 3\}$. If $T' \in \mathcal{H}$ and $T'(y) = H_{a,c}(y + e_i)$ for all $y \in W$ with $y_i \leq 1$, then $T' = H_{a,c-a_i}$.

(c) If $T \in \mathcal{H}$ and $\sigma \in \Pi$, then $T \circ \sigma_W \in \mathcal{H}$.

(d) In $y_{a,c}$, a cell p with $y_{a,c}(p) = 1$ and a direction $v \in D_3$ form a binary appearance if and only if $a \cdot v \not\equiv 0 \pmod{3}$; this holds for exactly 18 directions. If $y_{a,c}(m) = 0$, then there is $v \in \{\pm e_1, \pm e_2, \pm e_3\}$ such that $(m - v, v)$ is a binary appearance.

Proof. (a) $y_{a,c}(u + y) = [a \cdot y \equiv c - a \cdot u]$.

(b) Two different affine planes of $\mathbb{F}_3^3$ have at most three common points. The set $\{y : y_i \neq 2\}$ meets a plane in 9 points if the plane is $\{y_i = 0\}$ or $\{y_i = 1\}$, in no point if it is $\{y_i = 2\}$, and in 6 points otherwise. Hence two planes which have the same intersection with this set are equal. By (a) and the assumption, T' and $H_{a,c-a_i}$ agree on the cells with $y_i \leq 1$, that is, on this set.

(c) $H_{a,c}(\sigma_W(y)) = [\sigma^{-1}a \cdot (y - m_0) \equiv c - a \cdot m_0]$, because σ is orthogonal; this is $H_{a',c'}(y)$ with $a' = \sigma^{-1}a \in D_3$.

(d) The cells $p + v$ and $p + 2v$ have the residues $c + a \cdot v$ and $c + 2a \cdot v$; both differ from c if and only if $a \cdot v \not\equiv 0$. The map $v \mapsto a \cdot v$ is a surjective homomorphism $\{-1, 0, 1\}^3 = \mathbb{F}_3^3 \rightarrow \mathbb{F}_3$; its kernel has nine elements, so $26 - 8 = 18$ directions are not in the kernel. Let $y_{a,c}(m) = 0$ and $\rho := a \cdot m - c \not\equiv 0$. Choose i with $a_i \neq 0$ and $v = \pm e_i$ with $a \cdot v \equiv \rho$. Then $m - v$ has the residue c , and $m, m + v$ have the residues $c + \rho, c + 2\rho$, which are different from c . $\square$

Proof of Theorem 1.6(b) and (c). We use the second certificate of Proposition 6.1, with $\eta = \frac{47137}{294912}$.

(c) By Proposition 2.3 and Proposition 6.1(c),

$$48 \text{ ct}(\text{ABB}, \Gamma) \leq 48 N(x) = \sum_{\sigma \in \Pi} \sum_{u \in G} (\Psi_0 + Q_\lambda)(X^{\sigma,u}) \leq 48 \cdot 6|G| - \eta \#\{(\sigma, u) : X^{\sigma,u} \notin \mathcal{H}\}.$$

Since $\sigma(y) = \sigma_W(y) + \sigma(m_0) - m_0$, we have $X^{\sigma,u} = X^{u'} \circ \sigma_W$ with $u' = u + \sigma(m_0) - m_0$. By Lemma 6.2(c), $X^{\sigma,u} \in \mathcal{H}$ if and only if $X^{u'} \in \mathcal{H}$, and $u \mapsto u'$ is a bijection of G . So the number of pairs (σ, u) with $X^{\sigma,u} \notin \mathcal{H}$ is $48 \#\{u \in G : X^u \notin \mathcal{H}\}$, and (c) follows.

(b) Let $c_3(\text{ABB}, \Gamma) = 6$. By (c), or by Proposition 2.3, every binary appearance is an appearance and $X^u \in \mathcal{H}$ for all $u \in \mathbb{Z}^3$. Let $X^0 = H_{a,c_0}$. Since $X^{u+e_i}(y) = X^u(y + e_i)$, Lemma 6.2(b) and induction give $X^u = H_{a,c_0-a \cdot u}$ for all u with non-negative coordinates, hence

$$x(u) = X^u(0) = [a \cdot u \equiv c_0 \pmod{3}] \quad (u_1, u_2, u_3 \geq 0).$$

The periodicity of x under $n_i e_i$ forces $a_i n_i \equiv 0 \pmod{3}$, that is, $3 \mid n_i$ whenever $a_i \neq 0$. Then y_{a,c_0} has the same periods and agrees with x on a set of representatives, so $x = y_{a,c_0}$. If $x(m) = 0$, then by Lemma 6.2(d) the cell m is the second cell of a binary appearance $(m - v, v)$. This is an appearance, so the letter at m is B.

Conversely, let $3 \mid n_i$ whenever $a_i \neq 0$, and let Γ have the letter A at the cells with $a \cdot y \equiv c$ and B elsewhere. Then $y \mapsto a \cdot y$ is a surjective homomorphism $G \rightarrow \mathbb{Z}/3$, so the density of A is $\frac{1}{3}$, and by Lemma 6.2(d) each A-cell starts 18 appearances. Hence $c_3(\text{ABB}, \Gamma) = \frac{18}{3} = 6$. $\square$

The pairs (a, c) and $(-a, -c)$ give the same grid, and different planes give different grids. If exactly k sides are divisible by 3, there are $3^k - 1$ admissible vectors a , hence $\frac{3}{2}(3^k - 1)$ extremal grids of that shape, as stated after Theorem 1.6.

Proof of Corollary 1.7. As in the proof of Corollary 1.2(b), an $n \times n \times n$ array read as a grid of shape $(\mathbb{Z}/n)^3$ gives $M_3(n) \leq 6n^3$, for every $n \geq 1$. For the lower bound put the letter A at the cells y of the array with $y_1 \equiv 0 \pmod{3}$ and B elsewhere. A cell with the letter A and $2 \leq y_i \leq n-3$ for $i = 1, 2, 3$ starts an occurrence in each of the 18 directions with $v_1 \neq 0$. The inner cube $\{2, \dots, n-3\}^3$ consists of $(n-4)^2$ lines parallel to the first axis, and each of them contains at least $\lfloor (n-4)/3 \rfloor \geq (n-6)/3$ such cells. Hence $M_3(n) \geq 18(n-4)^2(n-6)/3 = 6(n-4)^2(n-6)$ for $n \geq 6$. $\square$

7 Higher dimensions

In this section w is an arbitrary word, and $N_v = N_v(w, \Gamma)$ denotes the number of appearances of w in Γ with direction v .

Lemma 7.1. $C_1(\text{AB}^L) = \frac{2}{L+1}$ for every $L \geq 1$. In particular $C_1(\text{ABB}) = \frac{2}{3}$ and $C_1(\text{ABBB}) = \frac{1}{2}$.

Proof. By Lemma 2.1 it suffices to bound the binary appearances of a function $x: \mathbb{Z}/n \rightarrow \{0, 1\}$; a constant function has none. Let $p_1, \dots, p_k$ be the points with $x = 1$ in cyclic order, and let g_i be the number of zeros between p_i and the next of these points. The binary appearance $(p_i, +1)$ exists if and only if $g_i \geq L$, and so does $(p_{i+1}, -1)$. Hence $N(x) = 2 \#\{i : g_i \geq L\}$, while $n = \sum_i (g_i + 1) \geq (L+1) \#\{i : g_i \geq L\}$. So $N(x) \leq \frac{2n}{L+1}$, with equality for $n = L+1$ and $k = 1$. $\square$

Lemma 7.2 (restriction to sublattices). *Let $u_1, \dots, u_k \in D_d$ have pairwise disjoint supports, and for $\varepsilon \in D_k$ put $v_\varepsilon := \sum_j \varepsilon_j u_j \in D_d$. Then every grid Γ of dimension d satisfies*

$$\sum_{\varepsilon \in D_k} N_{v_\varepsilon}(w, \Gamma) \leq C_k(w) |\Gamma|.$$

Proof. Let n be a common multiple of the sides of Γ . For $p \in G$ let Y_p be the grid of shape $(\mathbb{Z}/n)^k$ with $Y_p(i) := \Gamma(p + \sum_j i_j u_j)$. Then (i, ε) is an appearance of w in Y_p if and only if $(p + \sum_j i_j u_j, v_\varepsilon)$ is an appearance in Γ . For fixed i the map $p \mapsto p + \sum_j i_j u_j$ is a bijection of G . Hence

$$n^k \sum_{\varepsilon \in D_k} N_{v_\varepsilon}(w, \Gamma) = \sum_{p \in G} \text{ct}(w, Y_p) \leq |G| C_k(w) n^k. \quad \square$$

For $1 \leq j \leq d$ let

$$D_j := \frac{2}{\#\{v \in D_d : |v| = j\}} \sum_{|v|=j} \frac{N_v(w, \Gamma)}{|\Gamma|}$$

be twice the average of $N_v/|\Gamma|$ over the $\binom{d}{j} 2^j$ directions with j non-zero coordinates. Then $D_j \geq 0$ and

$$c_d(w, \Gamma) = \sum_{j=1}^d \binom{d}{j} 2^{j-1} D_j. \quad (11)$$

Apply Lemma 7.2 to $\sigma u_1, \dots, \sigma u_k$ for all $\sigma \in \Pi_d$ and average. The image of a fixed direction v under a uniformly random σ is uniformly distributed among the directions with $|v|$ non-zero coordinates. If

u_j has a_j non-zero coordinates, then v_ε has $\sum_{j:\varepsilon_j \neq 0} a_j$ of them. So, for all positive integers $a_1, \dots, a_k$ with sum at most d ,

$$\sum_{\emptyset \neq J \subseteq \{1, \dots, k\}} 2^{|J|-1} D_{a(J)} \leq C_k(w), \quad a(J) := \sum_{j \in J} a_j. \quad (12)$$

For $k = 1, 2, 3$ these are the inequalities

$$D_a \leq C_1, \quad D_a + D_b + 2D_{a+b} \leq C_2, \quad D_a + D_b + D_c + 2D_{a+b} + 2D_{a+c} + 2D_{b+c} + 4D_{a+b+c} \leq C_3.$$

Proof of Corollary 1.8. The lower bounds are $C_{d+1}(w) \geq 3C_d(w)$ [HS25, Lemma 2.2], applied to $C_3(\text{ABB}) = 6$ and $C_2(\text{ABBB}) = \frac{8}{3}$.

Let $w = \text{ABB}$, so that $C_1 = \frac{2}{3}$ and $C_3 = 6$ by Theorem 1.6. For $d = 4$ we use (12) with $(a_1, a_2, a_3) = (1, 1, 1)$ and $(1, 1, 2)$, and $D_3 \leq \frac{2}{3}$:

$$\begin{aligned} c_4 &= 4D_1 + 12D_2 + 16D_3 + 8D_4 \\ &\leq [3D_1 + 6D_2 + 4D_3] + 2[2D_1 + 3D_2 + 4D_3 + 4D_4] + 4D_3 \leq 6 + 12 + \frac{8}{3} = \frac{62}{3}. \end{aligned}$$

For $d = 5$ we use $(1, 1, 2)$ and $(1, 2, 2)$, and $D_3, D_4 \leq \frac{2}{3}$:

$$\begin{aligned} c_5 &= 5D_1 + 20D_2 + 40D_3 + 40D_4 + 16D_5 \\ &\leq 4[2D_1 + 3D_2 + 4D_3 + 4D_4] + 4[D_1 + 2D_2 + 4D_3 + 2D_4 + 4D_5] + 8D_3 + 16D_4 \\ &\leq 24 + 24 + \frac{16}{3} + \frac{32}{3} = 64. \end{aligned}$$

Let $w = \text{ABBB}$ and $d = 3$, so that $C_1 = \frac{1}{2}$ and $C_2 = \frac{8}{5}$ by Theorem 1.1. With $(a_1, a_2) = (1, 1)$ and $(1, 2)$ and $D_2 \leq \frac{1}{2}$,

$$c_3 = 3D_1 + 6D_2 + 4D_3 = \frac{1}{2}[2D_1 + 2D_2] + 2[D_1 + D_2 + 2D_3] + 3D_2 \leq \frac{4}{5} + \frac{16}{5} + \frac{3}{2} = \frac{11}{2}. \quad \square$$

Remark 7.3. The three bounds are the optimal values of the linear programs in the variables $D_j \geq 0$ given by (11) and by the inequalities (12) with $k \leq 3$ for ABB (where $C_2 = 2$) and $k \leq 2$ for ABBB ; for instance $D = (0, \frac{5}{9}, \frac{2}{3}, \frac{5}{12})$ satisfies all of them for $d = 4$ and has $c_4 = \frac{62}{3}$. For $d = 6$ one gets $C_6(\text{ABB}) \leq \frac{596}{3}$ in the same way, against the lower bound 162. The method cannot give $2 \cdot 3^{d-2}$ for $d \geq 4$: the constant sequence $D_j = \frac{6}{13}$ satisfies every inequality (12) with $k \leq 3$, and its value in (11) is $\frac{6}{13} \cdot \frac{3^d - 1}{2} > 2 \cdot 3^{d-2}$.

8 How the certificates were found

This section is not part of the proofs. The decimal numbers in it were computed in floating-point arithmetic, with the solver HiGHS [HH18] through SciPy [VGO⁺20]; they are *computed*, *not certified*, and nothing in the previous sections depends on them.

The relaxation. Let w be a word of length k on two letters, and let $\Psi_0(X)$ be the weighted number of arrows of the $k \times k$ window that read w in the pattern X , as in (4). The windows X^u of a grid of dimension 2, for a uniformly chosen $u \in G$, have a distribution whose marginals on the upper and on the lower $(k-1) \times k$ block are equal, and likewise on the left and on the right $k \times (k-1)$ block. Let $\text{LP}_k(w)$ be the maximum of $\sum_X \mu(X) \Psi_0(X)$ over all probability distributions μ with these two properties; then $C_2(\text{AB}^L) \leq \text{LP}_{L+1}(\text{AB}^L)$ by Lemmas 2.1 and 2.2. By linear programming duality, and by the symmetry of the problem, $\text{LP}_4(\text{ABBB})$ is the least M for which there is a function g with $\Psi_0 + D_g + D_g^t \leq M$. So the certificates of Sections 3 to 5 are optimal dual solutions of this linear program, and of its variants with a term for the density, with rational entries.

Values. Table 2 lists what was computed. For the words of length at most 4 on two letters the linear program on the window of the size of the word reproduces the known constants. For ABBB a window with a side shorter than 4 contains no diagonal arrow, so the 4×4 window is the smallest possible one; and if the potentials are restricted to 3×3 blocks inside this window, or to one of the two shift directions, the value is larger than $\frac{8}{5}$ (about $\frac{13}{8}$ and 2). With $M = \frac{8}{5}$ fixed, the largest uniform slack on the patterns outside $\mathcal{L}$ was $0.0428571 \approx \frac{3}{70}$, the constant of Theorem 1.3(b).

linear program	value	exact counterpart
LP_k for AB, ABA, ABB, AABB, ABAB, ABBA, AABA	3, 3, 2, $\frac{3}{2}$, 3, 2, 2	C_2 [AK24, HS25]
$LP_4(\text{ABBB})$	1.6000000	Proposition 3.1
potentials on 3×3 blocks only	1.6250	none
one shift direction only	2.0000	none
method of [HS25, §5.3], 5×5 window	1.678728	$\frac{59526}{35459}$ [HS25]
ABBB, density α of A prescribed	as in Theorem 1.4(a)	Proposition 5.1
ABB, $d = 3$, polynomials of degree ≤ 2	6.496170	none
ABB, $d = 3$, polynomials of degree ≤ 3	6.047479	none
ABB, $d = 3$, polynomials of degree ≤ 4	6.000000	Proposition 6.1

Table 2: Values of linear programs, computed in floating-point arithmetic and not certified, except where an exact counterpart is named. The line for the method of [HS25] was computed with a program of this note which implements that method.

Rounding. A dual solution in the relative interior of the optimal face was computed, its entries were rounded to multiples of $1/q$, and the finitely many forced equalities were restored by solving a small linear system exactly. The result was then checked in integers on all patterns; whatever passes this check is a certificate, however it was found. For (6) the denominator $q = 20$ already works, and $q = 280$ keeps the slack $\frac{3}{70}$; the second verification run used $q = 2520$.

Dimension three. The window $\{0, 1, 2\}^3$ has 2^{27} patterns. The potentials were polynomials Q_λ of bounded degree which are invariant under the symmetries of the cube and satisfy the hypothesis of Lemma 2.5; the inequalities were generated by cutting planes, with an evaluation of all 2^{27} patterns in every round. Degrees 2 and 3 are not sufficient (Table 2); degree 4 gives 6. The first certificate was obtained by rounding a solution to multiples of $\frac{1}{10080}$ and solving the three independent equalities on the hyperplane windows exactly. The second verification run maximised the uniform slack outside $\mathcal{H}$, found 0.1602, and rounded to multiples of 2^{-16} ; this gave the constant $\frac{47137}{294912} = 0.1598\dots$ of Theorem 1.6(c). A small example of the same kind, for which Proposition 2.3 has only 512 patterns to consider: for ABB in dimension 2 the polynomial Q on the window $\{0, 1, 2\}^2$ with the coefficients $-\frac{2}{3}$ at the corners, $\frac{1}{3}$ at the midpoints of the edges, $\frac{4}{3}$ at the centre m , $\frac{1}{3}$ at the pairs $\{m, \text{corner}\}$ and $-\frac{1}{3}$ at the pairs of midpoints of adjacent edges satisfies the hypothesis of Lemma 2.5 and $\Psi_0 + Q \leq 2$ (program `check_abb2d.py`); this gives another proof of $C_2(\text{ABB}) \leq 2$ [HS25, Lemma 5.4].

9 Computations and reproducibility

The reproducibility package contains the eight certificates (in `certificates/`; none is larger than 60 kilobytes), all programs with their recorded outputs, and a script `run_all.sh` which runs the checks of the first four sources below again, one process at a time, and compares the results with the recorded ones (the six largest enumerations, the simulated annealing and the floating-point linear programs

are not part of the script); the programs of the fifth source have a script of their own, `run_run2.sh`. The programs come from five sources.

The original computation (original/). The checkers of the note: `check_gcert.py` (Proposition 3.1; pure Python) and a second implementation with bit masks and 64-bit integers, also for (7); `check_extremal.py` (Proposition 4.1 and Lemma 4.2(b),(c)); `check_density.py` ((7) and the grids of Theorem 1.4); `check_abb3d.py` (Proposition 6.1 and Lemma 6.2(b)); `check_abb2d.py` (the example at the end of Section 8). Further: tests of the identity in (5) with the certificates on random grids, also with sides shorter than the window and with a third letter; tests that corrupted certificates are rejected; the programs for Section 8; and enumerations of all grids on two letters on all tori with at most 25 cells for $d = 2$ and at most 24 cells for $d = 3$.

The first verification run (verification_run_A/). A checker for dimension 2 in pure Python with exact fractions. For dimension 3, a program which expands a certificate and tests Proposition 6.1(a), and two C programs for Proposition 6.1(b),(c). One uses the subset-sum transform and computes Ψ_0 from the 49 lines. The other walks through all patterns in the order of a Gray code, updates Q_λ incrementally without a table, and computes Ψ_0 from the 98 arrows. Both print a checksum of the values at all 2^{27} patterns, and the checksums agree. Further: rejection tests; the identities of Lemmas 2.2, 2.4 and 2.5 with the certificates on 262 grids; Proposition 5.3; examples for Corollary 1.2(b), Corollary 1.7 and Corollary 1.8; an enumeration of all grids on two letters on all tori with at most 30 cells and on 2×16 , 4×8 , 5×7 , 6×6 for $d = 2$, and on all tori with at most 28 cells and on $2 \times 3 \times 5$, $2 \times 4 \times 4$, $2 \times 2 \times 8$ for $d = 3$; and simulated annealing on 52 larger tori.

The second verification run (verification_run_B/). Its linear programs, the programs that produced its four certificates, and its exact checkers `verify2d.py` (pure Python) and `verify3d.py` (NumPy [HMvdW⁺20]). Further: Lemma 4.2; an enumeration of all grids on two letters on all tori with at most 25 cells for $d = 2$ (160,098,126 grids) and at most 24 cells for $d = 3$ (127,924,302 grids); the linear programs of Remark 7.3 with rational solutions; and simulated annealing.

Programs written with this note (writing_stage/). A further exact checker for (6), (7) and (8), in pure Python; the count of the grids of Proposition 5.3 from the definition; and tests of the vertices of Q , of Lemma 5.2 (on the sides of Q and on 200 pairs of random grids), of Theorem 1.4(a), and of the inequalities in the proof of Corollary 1.5(c) (on all 111,954 grids with three letters of ten small shapes and on 700 further grids).

The third verification run (independent_run_2/: the second round of verification). A checker for dimension 2 in exact fractions, with Ψ_0 from the 20 arrows of (4). For dimension 3, a program for Proposition 6.1(a) and a C program for Proposition 6.1(b),(c) which uses neither a table of 2^{27} numbers nor a Gray code (for each of the 2^{13} patterns of thirteen cells, a subset-sum transform over the other fourteen; Ψ_0 from the 98 arrows); it is compared with a direct evaluation in fractions at 1694 patterns and with the exact sum of all 2^{27} values. Further, from the definition: the grids of Theorems 1.1, 1.3, 1.4 and 1.6 and the 190 grids of Proposition 5.3 with $N \leq 400$; Lemma 5.2 on 400 pairs of random grids; the proof of Corollary 1.5(c), cell by cell, on 116,590 grids with three or four letters; (5) with all eight certificates on 150 grids; (12) on 170 grids and the linear programs of Remark 7.3, solved exactly; an enumeration of all grids on two letters on the 58 tori with at most 30 cells ($d = 2$) and on the 62 tori with at most 28 cells ($d = 3$), with a test of Theorem 1.3(b), Theorem 1.4(a) and Theorem 1.6(c) for every grid; and simulated annealing on 28 larger tori.

Results. Every certificate was accepted by every program that can read it: (6) for its three certificates and (7) for its two certificates by six programs each (two of the note, one of each of the three runs, one written with the note), (8) by four programs, and Proposition 6.1 for each of its two certificates by five programs (of the note, of the second run, the two C programs of the first run, and the C program of the third run). No search found a grid with concentration above $\frac{8}{5}$, respectively above 6. Among all shapes enumerated, the value $\frac{8}{5}$ occurs only on the torus 5×5 , exactly for the

ten lattice grids, and the value 6 occurs exactly on the shapes with a side divisible by 3, for $\frac{3}{2}(3^k - 1)$ grids, as Theorem 1.6(b) says. The inequalities of Theorem 1.3(b) and of Theorem 1.4(a) hold for all grids on two letters on the tori with at most 30 cells, and that of Theorem 1.6(c) on the tori with at most 28 cells.

Re-runs. On 10 October 2026 the script `run_all.sh` was run from the package with all its parts, one process at a time, on a laptop (Apple M3 Max; Python 3.13.5, NumPy 2.5.2, SciPy 1.18.1, Apple clang 21). All 72 steps ended without error, and the outputs agree with the recorded ones. Each check in dimension 2 takes at most two seconds. In dimension 3, per certificate, `check_abb3d.py` takes about 16 seconds and 1.3 gigabytes, `verify3d.py` about 16 seconds and 2.5 gigabytes, the C program with the subset-sum transform 14 seconds and 1.1 gigabytes, and the Gray-code walk 83 and 169 seconds for the two certificates. The enumerations take about 15 minutes; the six largest shapes of the first run, with 32 to 36 cells, were enumerated again separately (84 minutes), with the recorded maxima and numbers of maximisers. The C program of the third run takes 26 seconds, in one process and with two megabytes of memory. The third run ran `run_all.sh` once more with all its parts, from an extracted copy of the archive, with the same result (72 steps without error, 73 output files as recorded), and then its own script in the same way (21 output files as recorded).

Verification. The results were first obtained, with proofs, with the four certificates of the note and with the programs in `original/`, on 10 October 2026. Three independent verification runs, all AI-assisted, followed on the same day; each wrote its own programs before reading those of the note. The first two examined the first written version of the results, and neither saw the report of the other. The third examined the final text of this note and had the reports of the first two.

The first run examined the statements and the proofs. It compared the definitions and the questions with the source of [HS25] and with pages 71–74 of [FKS26]; re-derived Lemmas 2.1, 2.2, 2.4 and 2.5, the deduction of Theorems 1.1, 1.3, 1.4 and 1.6 from the finite inequalities, Corollaries 1.2 and 1.8 and the refutation in Corollary 1.5(b), also for grids with sides shorter than the window; and checked every finite statement with the programs described above. It found no error in a theorem. It required three corrections of the first written version: two inaccurate remarks about the sources (the two words of Question 11 are ordinary text in the report and are present in the corpus record; a remark on counting lines without orientation was deleted), and the addition of Corollary 1.7, without which the case $d = 3$ of Question 11 would have been proved only in the normalisation of [HS25]. It also supplied the family of Proposition 5.3, with the argument by locality, and the correct ranges of n in Corollary 1.2(b).

The second run recomputed the results without the certificates of the note. Its own linear program gave the value 1.6000 on the 4×4 window (in floating-point arithmetic: computed, not certified), and it produced and verified, in exact arithmetic, its own certificates for (6), for (7), for the new inequality (8) and for Proposition 6.1, the last with the larger constant of Theorem 1.6(c). Its checkers accept the certificates of the note, its enumerations agree with those of the note in all 97 shapes, and it re-derived Lemma 2.5, Lemma 7.2 and Corollary 1.8. It found no error in a theorem and required the correction of the first of the two remarks, the addition of Corollary 1.7, a more careful statement about the citation databases, the label “computed, not certified” for the values of Section 8, and the credits given below. All corrections required by the two runs were made.

When this note was written, the following was added to what the two runs had examined. Theorem 1.4 was completed by the third bound, with the certificate of the second run, and by part (b) with Lemma 5.2, which generalises the argument of the first run for Proposition 5.3; the formula of Proposition 5.3 is stated for all admissible L, N . Theorem 1.6(c) is stated with the constant of the second certificate. These two certificates of the second run were checked again with the programs of the note and of the first run. The bound for arbitrary alphabets in Corollary 1.5(c) was added, and

the proofs of Lemma 4.2(b), Lemma 6.2(b) and Lemma 7.2 were written in their present form. The programs in `writing_stage/` test these additions.

The third run examined the final text, and these additions first. It re-derived Lemma 5.2 with its boundary terms, also when a part has fewer than seven columns; both inclusions of Theorem 1.4(b); Proposition 5.3; the bound for arbitrary alphabets in Corollary 1.5(c); the proofs of Lemma 4.2(b), Lemma 6.2(b) and Lemma 7.2; and then all other proofs, line by line. It compared the questions, the definitions and every numbered citation with pages 72–73 of [FKS26], including the figures, and with the source of version 2 of [HS25], and it repeated the literature search. It found no wrong theorem, proposition or lemma and no gap. It required twelve corrections of the text, with the corresponding changes of the package, which were made; none changes the content of a statement. Among them: the theorem of [AK24] had been understated (Section 1.2); the weights of [HS25, Proposition 5.2] had been called non-negative (Section 1.4); one sentence was added to the proof of Lemma 4.2(b); and the floating-point value above was labelled. Its programs, described above, reproduce every finite statement of this note: the numbers in Propositions 3.1, 4.1, 5.1 and 6.1, all counts of grids, and the lists of the extremal grids on the tori which it enumerated. Its maxima agree with those of the two earlier runs on all tori which they share (120 and 97).

Scope and priority. The note determines $C_2(\text{ABBB})$ (Question 9.3 of [HS25]; the last part of Question 11 of [FKS26]) and $C_3(\text{ABB})$ (the case $d = 3$ of the first part of Question 11 and of Question 9.2 of [HS25]), and it classifies the extremal grids in these two cases, which settles Question 10 of [FKS26] (Question 2.5 of [HS25]) for these two pairs only. Question 10 in general and the values $C_d(\text{ABB})$ for $d \geq 4$ stay open; for $d = 4, 5$ only the bounds of Corollary 1.8 are proved. The two main theorems are computer-assisted in the sense of Section 1.4: the reductions are proved in the text, and the finite inequalities are verified in exact integer arithmetic by programs which anyone can run again in seconds or minutes.

The problem goes back to Patchell and Spiro [PS22] and to Alon and Kravitz [AK24]. The setting of this note, the bounds (1) with the notion of stackability, the five-queens grid with the lower bound $\frac{8}{5}$ and the conjectured value, the value $C_2(\text{ABB}) = 2$, the bound for ABB on grids of shape $(\mathbb{Z}/3)^d$, the local linear-programming method with the bound $\frac{59526}{35459}$, and Questions 2.5, 9.2, 9.3, 9.5 and 9.6 are due to Halberstam and Schildkraut [HS25]; Questions 10 and 11 are from Schildkraut’s contribution to the Oberwolfach problem session [FKS26]. The lower bounds in Theorems 1.1 and 1.6 and Corollary 1.8 are theirs. The technique of the upper bounds is standard (Section 1.4).

What was read: pages 72–73 of [FKS26] in the publisher’s file, with the figures; of [HS25], in version 2, to which all our numbers refer, Sections 1, 2, 5.1–5.4, 6.1 and 9, the passages of Section 6.2 on ABBBB and BABBB , the statements of Propositions 3.1, 3.2, 8.1 and of Lemma 3.9, and the passage of Appendix A with the bound for ABBB ; its proofs were not checked, except that of Lemma 2.2, which is the only result of [HS25] used in our proofs. We did not read [PS22, AK24]; what we say about them is taken from [HS25, FKS26]. Of [Sto14, HKD⁺16] only the abstracts were read, and [Raz07] is cited as in [HS25, §2.5].

The searches were made on 10 October 2026: the versions of [HS25] on arXiv (version 1 of 24 November 2025, version 2 of 26 November 2025, no later version) and the arXiv listings of its authors; arXiv queries for the title words, for word searches in grids, queens, stackable words, ABBB , and for [PS22, AK24]; the lists of citing works in OpenAlex and in Semantic Scholar; Crossref, zbMATH Open, DataCite and Zenodo (queries for the constants, the title, the authors and the record identifier); the publisher’s page of the report; and six web searches. The third verification run repeated the queries to these databases later on the same day and made one further web search. We found no proof of either value, no classification of the extremal grids, and no later version of [HS25] announcing one. Semantic Scholar lists no later work citing [HS25], [AK24] or [PS22]; OpenAlex lists none either, but

its citation data are incomplete for these papers (it shows no citations of [PS22, AK24], although [HS25] cites both). MathSciNet and Google Scholar were not consulted. A search that finds nothing is not a proof of novelty, and we make no claim of priority. The sources are recent and their authors are active; this note was written without knowledge of unpublished progress of the authors of the problem.

AI-assisted tools were used to search the problem corpus and the literature, to find the arguments, to write and run the verification programs, and to help prepare this manuscript. The author has reviewed the proofs, the computations and the sources and is responsible for all claims. The note has not been peer reviewed.

References

- [AK24] Noga Alon and Noah Kravitz. Cats in cubes. *The Electronic Journal of Combinatorics*, 31(3):Paper No. P3.29, 2024. [arXiv:2211.14887](#), [doi:10.37236/11735](#).
- [FKS26] Jacob Fox, Peter Keevash, and Wojciech Samotij. Combinatorics. *Oberwolfach Reports*, 23(1):5–79, 2026. Report No. 1/2026 (the named authors are the organisers). Problem session: C. Schildkraut, pp. 72–73. [doi:10.4171/OWR/2026/1](#).
- [HH18] Qi Huangfu and J. A. Julian Hall. Parallelizing the dual revised simplex method. *Mathematical Programming Computation*, 10(1):119–142, 2018. [doi:10.1007/s12532-017-0130-5](#).
- [HKD⁺16] Wenxuan Huang, Daniil A. Kitchaev, Stephen T. Dacek, et al. Finding and proving the exact ground state of a generalized Ising model by convex optimization and MAX-SAT. *Physical Review B*, 94(13):134424, 2016. [arXiv:1604.06722](#), [doi:10.1103/PhysRevB.94.134424](#).
- [HMvdW⁺20] Charles R. Harris, K. Jarrod Millman, Stéfan J. van der Walt, et al. Array programming with NumPy. *Nature*, 585(7825):357–362, 2020. [doi:10.1038/s41586-020-2649-2](#).
- [HS25] Zachary Halberstam and Carl Schildkraut. Words with repeated letters in a grid. Preprint, 2025. Version 2 of 26 November 2025; 42 pages. [arXiv:2511.19678](#).
- [PS22] Gregory Patchell and Sam Spiro. The maximum number of appearances of a word in a grid. *The American Mathematical Monthly*, 129(5):415–434, 2022. [arXiv:2207.11273](#), [doi:10.1080/00029890.2022.2040319](#).
- [Raz07] Alexander A. Razborov. Flag algebras. *The Journal of Symbolic Logic*, 72(4):1239–1282, 2007. [doi:10.2178/jsl/1203350785](#).
- [Sto14] Derrick Stolee. Automated discharging arguments for density problems in grids. Preprint (extended abstract), 2014. [arXiv:1409.5922](#).
- [ula26] ulamai. UnsolvedMath. Hugging Face dataset, <https://huggingface.co/datasets/ulamai/UnsolvedMath>, 2026. Version 1.6.0; record OWR-14299905-036.
- [VGO⁺20] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, 17(3):261–272, 2020. [doi:10.1038/s41592-019-0686-2](#).