

The Maximum Number of Appearances of a Word in a Grid

Gregory Patchell and Sam Spiro

Abstract. How can you fill a 3×3 grid with the letters A and M so that the word “AMM” appears as many times as possible in the grid? More generally, given a word w of length n , how can you fill an $n \times n$ grid so that w appears as many times as possible? We solve this problem exactly for several families of words, and we asymptotically solve this problem in higher-dimensional grids.

1. INTRODUCTION. Consider the following problem. You have been asked to advertise *The American Mathematical Monthly* using a 3×3 grid. Specifically, you are asked to fill in the squares of the grid with A’s and M’s to maximize the number of times the word “AMM” appears along a horizontal, vertical, or diagonal line; and we will also count a line if it includes the word written backwards.

Given this task, you might think for a bit and come up with a number of designs that seem to do well. Maybe your first thought is to fill in the squares as in Figure 1, which has 5 instances of the word “AMM.”

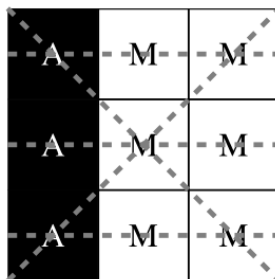


Figure 1. A 3×3 grid with 5 instances of the word AMM.

However, being the mathematician that you are, you are not convinced that this is the best you can do. You might think for a bit and come up with more complicated designs like those in Figure 2, which also all contain 5 copies of the word. Is there a design giving more than 5 copies of the word AMM?

Of course, once we successfully solve this problem, other organizations are going to want to hire us for similar advertising jobs. To prepare for these new jobs, we let $[n] = \{1, 2, \dots, n\}$ and $[n]^2 = \{(i, j) : i, j \in [n]\}$. We define an n -grid G to be a function from $[n]^2$ to a set of letters, and we will simply call this a *grid* whenever n is understood. Given a word $w = w_1 \cdots w_n$ and an n -grid G , we say that the i th row of G contains w if $G(i, j) = w_j$ for all $1 \leq j \leq n$, or if $G(i, j) = w_{n-j+1}$ for all $1 \leq j \leq n$. We similarly define what it means for the i th column of G to contain w , as well as for the diagonals of G to contain w . We let $f(w, G)$ be the total number of rows, columns, and diagonals of G containing w , and we define $f(w) = \max_G f(w, G)$, where the maximum ranges over all n -grids G . More generally, if \mathcal{W} is a set of words of length n , we define $f(\mathcal{W}, G)$ to be the number of rows, columns, and diagonals of G containing some $w \in \mathcal{W}$, and we similarly define $f(\mathcal{W})$.

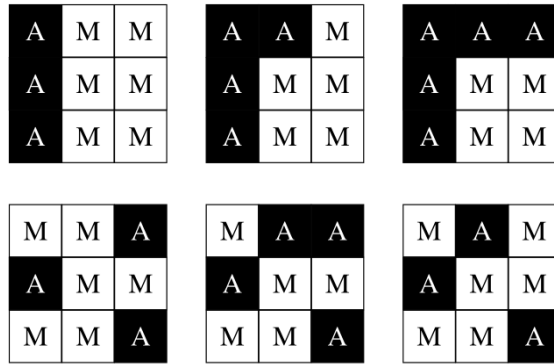


Figure 2. Several grids with 5 instances of AMM.

We note that determining which grids G satisfy $f(\mathcal{W}, G) = f(\mathcal{W})$ can be seen as a generalization of a classical problem concerning Latin squares. In our language, a *Latin square* is an n -grid G such that every row and column of G contains a permutation of the elements of $[n]$; see Figure 3 for an example.

1	2	3	4
4	3	2	1
2	1	4	3
3	4	1	2

Figure 3. A (diagonal) Latin square.

If both diagonals of G also contain a permutation on $[n]$ (as is the case in Figure 3), we call G a *diagonal Latin square*. Observe that G being a diagonal Latin square is equivalent to having $f(\mathcal{P}, G) = 2n + 2$ where \mathcal{P} is the set of all permutations of $[n]$. Thus grids G with $f(\mathcal{W}, G) = f(\mathcal{W})$ where \mathcal{W} is a set of words can be viewed as a generalization of diagonal Latin squares. Much more can be said about Latin squares; see, for example, [4, 5, 6, 8, 10, 11].

In this paper we focus primarily on the problem of determining $f(w)$ for a single word w , and even this humble task seems to be difficult to solve in general. Nevertheless, we are able to compute $f(w)$ exactly for several natural choices of words w .

As a point of reference, for all words w of length n , we will show that

$$n + 2 \leq f(w) \leq 2n + 2.$$

This lower bound tends to be closer to the truth when w has no letter appearing many times, or when w is symmetric.

Theorem 1. *If w is a word of length $n \geq 2$ such that each letter occurs at most $n/4$ times, then*

$$f(w) = n + 2.$$

Moreover, for infinitely many $n \geq 2$ there exists a word w of length n such that every letter occurs at most $1 + n/4$ times and such that $f(w) > n + 2$.

Theorem 2. *If w is a word of length $n \geq 2$ such that $w_i = w_{n-i+1}$ for all i , then*

$$f(w) = \max\{n, 2k\} + 2,$$

where k is the maximum number of times any letter appears in w .

Conversely, when a word is “anti-symmetric,” the upper bound tends to be closer to the truth.

Proposition 1. *If w is a word of length $n \geq 2$ using only two letters, and if $w_i \neq w_{n-i+1}$ for all i , then*

$$f(w) = 2n.$$

This together with Theorem 2 gives the following peculiar result, which demonstrates how sensitive $f(w)$ can be to the symmetries of w .

Corollary 1. *Let w be the word of length $n \geq 2$ defined by $w = \text{AMAM} \cdots$. Then*

$$f(w) = \begin{cases} n + 2 & n \text{ odd,} \\ 2n & n \text{ even.} \end{cases}$$

And of course, we also compute $f(\text{AMM})$. More generally, we prove the following, where by $A^k M^{n-k}$ we mean the word which starts with k copies of the letter A followed by $n - k$ copies of the letter M.

Theorem 3. *Let $w = A^k M^{n-k}$ with $1 \leq k \leq n/2$ and $n \geq 2$. Then*

$$f(w) = \max\{2(n - k) + 1, 4k\}.$$

In particular, taking $n = 3, k = 1$ shows that indeed $f(\text{AMM}) = 5$.

We consider a generalization of this problem to higher-dimensional grids, which might be useful, for example, if the AMM wanted to expand its readership to 4-dimensional beings. We formally make our definitions in Section 3. Informally, we let $f(w, d)$ be the maximum number of lines a d -dimensional n -grid can have which contain w . We recall the notation $f \sim g$ for functions f, g to mean that $\lim_{d \rightarrow \infty} \frac{f(d)}{g(d)} = 1$. With this in mind, we have the following.

Theorem 4. *Let w be a word of length $n \geq 2$. If $w_i = w_{n-i+1}$ for all i , then*

$$f(w, d) \sim \frac{1}{2}(n + 2)^d,$$

and otherwise

$$f(w, d) \sim \frac{1}{4}(n + 2)^d.$$

As a point of reference, it is well-known that d -dimensional n -grid contains roughly $\frac{1}{2}(n + 2)^d$ total lines [3]. Thus, despite the fact that $f(w, 2)$ can be quite small for

symmetric words, in higher-dimensional grids we can make it so that almost every line contains a copy of w .

Determining $f(w, d)$ exactly for all d seems very difficult in general. Remarkably, we can do this for certain anti-symmetric words, namely those mentioned in Proposition 1.

Theorem 5. *If w is a word of length $n \geq 2$ using only two letters, and such that $w_i \neq w_{n-i+1}$ for all i , then for all d ,*

$$f(w, d) = \frac{1}{4}((n + 2)^d - (n - 2)^d).$$

Another natural generalization of this problem is to bound the maximum number of times a word w of length k can appear in an (n, d) -grid with $k \leq n$, which we denote by $f(w, n, d)$. Even in the $d = 2$ case and for simple words we do not have asymptotic formulas for $f(w, n, d)$, though we are able to obtain some reasonable asymptotic bounds. We briefly discuss these results in Section 4.

2. WORDS ON A PLANE.

General lower bounds. From now on, whenever we have a word w , we assume it is of length n and that $n \geq 2$. We first consider basic constructions of grids to give lower bounds on $f(w)$. The simplest such result is the following.

Lemma 1. *For all w ,*

$$f(w) \geq n + 2.$$

For example, Figure 4 demonstrates how we can achieve 5 lines for $w = ABC$.

A	B	C
A	B	C
A	B	C

Figure 4. A 3×3 grid with 5 instances of the word ABC.

Proof. Let G be the grid that has $G(i, j) = w_j$ for all i, j . Then every row and diagonal of G contains w , so $f(w) \geq f(w, G) \geq n + 2$. ■

We can improve upon this bound by exploiting certain structures of the given word w . For example, we have the following.

Lemma 2. *Let w be a word such that some letter A appears k times. Then*

$$f(w) \geq 2k + 1.$$

Further, if $w_i = w_{n-i+1}$ whenever $w_i = A$, then

$$f(w) \geq 2k + 2.$$

Figure 5(a) gives an example of the construction of Lemma 2 with the word BAACA and Figure 5(b) does the same with the word ABACA. Note that the latter grid achieves its target word ABACA in both of its diagonals while the former does not.

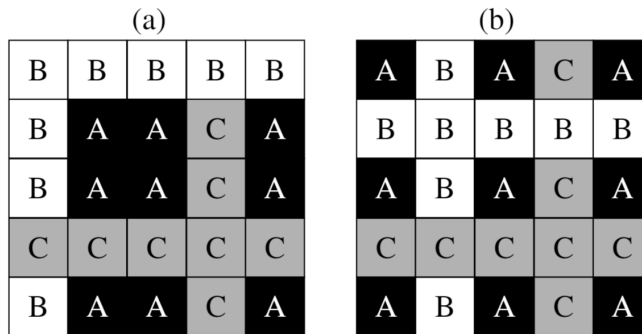


Figure 5. (a) A grid with 7 copies of BAACA. (b) A grid with 8 copies of ABACA.

Proof. Let I be the set of indices with $w_i = A$ (this corresponds to $I = \{2, 3, 5\}$ in Figure 5(a) and $I = \{1, 3, 5\}$ in Figure 5(b)). Define the grid G by setting $G(i, j) = w_j$ whenever $i \in I$ and $G(i, j) = w_i$ otherwise. That is, we write out w along each row corresponding to I , and we then try to do this along each column. By construction, the i th row of G contains w whenever $i \in I$, and also $G(i, i) = w_i$ for all i . If $i \in I$, we claim that the i th column also contains w , that is, $G(j, i) = w_j$ for all j . This is immediate if $j \notin I$; otherwise $G(j, i) = w_i = A = w_j$. Thus in total we have that $f(w) \geq f(w, G) \geq 2|I| + 1 = 2k + 1$.

If $w_i = w_{n-i+1}$ for all $i \in I$, then we claim that $G(i, n - i + 1) = w_i$ for all i . This is immediate if $i \notin I$, and otherwise $G(i, n - i + 1) = w_{n-i+1} = A = w_i$. This gives an extra diagonal containing w , proving the second bound. ■

While Lemma 2 gives a small gain when w is somewhat symmetric, we can often do much better if the word is anti-symmetric.

Lemma 3. *Let w be a word and A, M two letters used in w . Define*

$$T = \{i : w_i = A, w_{n-i+1} = M\},$$

and let $t = |T|$. Then

$$f(w) \geq 4t.$$

Moreover, if w only uses the letters A and M , then

$$f(w) \geq n + t.$$

Figure 6(a) gives the $4t$ construction and Figure 6(b) gives the $n + t$ construction for the word AMAAM. In this example, $T = \{1, 4\}$ and X is an arbitrary letter.

Proof. Define $S = \{i : w_{n-i+1} = A, w_i = M\}$, and note that by definition this set is disjoint from T . Furthermore, $|S| = t$ since $i \in T$ if and only if $n - i + 1 \in S$. Let

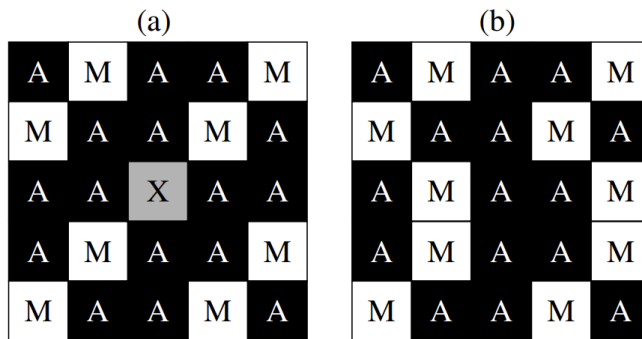


Figure 6. (a) A grid with $4 \cdot 2$ copies of AMAAM. (b) A grid with $5 + 2$ copies of AMAAM.

G be the grid with

$$G(i, j) = \begin{cases} w_j & i \in T, \\ w_i & j \in T, \\ w_{n-j+1} & i \in S, \\ w_{n-i+1} & j \in S, \\ X & \text{otherwise,} \end{cases}$$

where X is an arbitrary letter. It is not difficult to see that this grid is well-defined when either $i, j \in T$ or $i, j \in S$ (since $w_i = w_j$ and $w_{n-i+1} = w_{n-j+1}$ in these cases). If, say, $i \in T$ and $j \in S$, then $w_{n-i+1} = M = w_j$, so it is well-defined in this case as well. By construction, every row and column corresponding to T and S contains w , so $f(w) \geq f(w, G) \geq 4t$.

If w consists of only A's and M's, then we define the grid G by $G(i, j) = w_j$ if $w_i = A$ and $G(i, j) = w_{n-j+1}$ otherwise. By construction, all of the n rows of G contain w written forwards or backwards. We also claim for $i \in T$ that the i th column of G contains w . Indeed, for such i , if $w_j = A$ then $G(j, i) = w_i = A = w_j$, and otherwise $G(j, i) = w_{n-i+1} = M = w_j$. Thus for every $i \in T$, $G(j, i) = w_j$, so each of these columns contains w and we have $f(w) \geq f(w, G) \geq n + t$. ■

General upper bounds. Because there are $2n + 2$ lines in an n -grid, we have $f(w) \leq 2n + 2$ for all w . This bound is sharp when w is the constant word A^n . We can slightly improve this bound for nonsymmetric words.

Lemma 4. *Let w be a word such that $w_i \neq w_{n-i+1}$ for some i . Then*

$$f(w) \leq 2n.$$

Proof. Let G be an n -grid and define

$$Q = \{(i, i), (i, n - i + 1), (n - i + 1, i), (n - i + 1, n - i + 1)\},$$

$$Q^+ = \{s \in Q : G(s) = w_i\}, \quad Q^- = \{s \in Q : G(s) = w_{n-i+1}\}.$$

Given two distinct $q, q' \in Q$, let $\ell_{q,q'}$ be the unique line of G containing q and q' . Observe that this line can contain w only if either $q \in Q^+$ and $q' \in Q^-$ or the other

way around. Thus out of the 6 lines $\ell_{q,q'}$, the number of these that can contain w is at most

$$|Q^+||Q^-| \leq \left(\frac{1}{2}|Q|\right)^2 = 4.$$

In particular, at least two of the $2n + 2$ lines of G cannot contain w , giving the result. ■

Next we consider an upper bound that works well when there are not too many copies of a given letter.

Lemma 5. *Let w be a word such that each letter appears at most k times in w . Then*

$$f(w) \leq \max\{4k, n\} + 2.$$

Proof. Let G be a grid and assume $f(w, G) > 4k + 2$. Let R_1 denote the set of $i \in [n]$ such that $G(i, j) = w_j$ for all j , R_2 the set of $i \in [n]$ such that $G(i, j) = w_{n-j+1}$ for all j , and similarly define C_1 and C_2 for the columns. Note that $f(w, G) \leq |R_1| + |R_2| + |C_1| + |C_2| + 2$, and because $f(w, G) > 4k + 2$, one of these sets must have size at least $k + 1$. Without loss of generality we can assume that R_1 has this property.

We claim that no column of G contains w . Indeed, in the i th column we have at least $k + 1$ different $j \in R_1$ such that $G(j, i) = w_i$. However, each letter of w appears at most k times, so this column cannot contain w . Thus for such a grid we have $f(w, G) \leq n + 2$. ■

A similar proof gives strong bounds when a word is very symmetric.

Lemma 6. *Let w be a word such that each letter appears at most k times in w and such that there are s indices i with $w_i = w_{n-i+1}$. Then*

$$f(w) \leq \max\{n + 2k - s, n\} + 2.$$

Proof. Let G be a grid and define R_1, R_2, C_1, C_2 as in the proof of Lemma 5. We can assume without loss of generality that $|R_1| + |R_2| \geq |C_1| + |C_2|$. If $|R_1|$ or $|R_2|$ is strictly larger than k , then no column of G can contain w (since each column would contain more than k copies of some letter), so $f(w) \leq n + 2$ in this case. Thus we can assume $|R_1|, |R_2| \leq k$.

If $|R_1| + |R_2| > k$ and i is such that $w_i = w_{n-i+1}$, then the i th column of G contains more than k copies of the letter w_i and hence cannot contain w . Because there are s such i , we have

$$f(w, G) \leq |R_1| + |R_2| + |C_1| + |C_2| + 2 \leq 2k + n - s + 2.$$

Finally, if $|R_1| + |R_2| \leq k$ then also $|C_1| + |C_2| \leq k$ by assumption, and we have $f(w, G) \leq 2k + 2 \leq 2k + n - s + 2$, where this last step used $s \leq n$. ■

We claim without proof that the argument used to prove Lemma 5 generalizes as follows, which could be useful for bounding $f(w)$ for other kinds of words.

Proposition 2. *Let w be a word on the letters A_1, \dots, A_m such that A_i appears k_i times in w with $k_1 \geq k_2 \geq \dots \geq k_m$. Then for all grids G and all $1 \leq i \leq m$,*

$$f(w) \leq \max\{4k_i, n + \sum_{j < i} k_j\} + 2.$$

Note that the $i = 1$ case is exactly the statement of Lemma 5, and for example the $i = 2$ case gives better bounds for $f(A^{n/3}M^{2n/3})$. A similar generalization can be made for Lemma 6.

Putting the pieces together. In this section we use the lemmas we have developed to prove our main results for $f(w)$. The first few results are immediate.

Proof of Theorem 1. For the first result, the lower bound follows from Lemma 1 and the upper bound from Lemma 5. For the second result, take $n = 4k$ for any integer $k \geq 1$. Consider the word w that begins with $k + 1$ copies of the letter A , that ends with $k + 1$ copies of the letter M , and that has $2k - 2$ distinct letters in the remaining positions. Note that each letter in w appears at most $k + 1 = n/4 + 1$ times. By Lemma 3, we have

$$f(w) \geq 4(k + 1) > 4k + 2 = n + 2. \quad \blacksquare$$

Proof of Theorem 2. The lower bounds follow from Lemmas 1 and 2. The upper bound follows from Lemma 6. \blacksquare

Proof of Proposition 1. The lower bound follows from Lemma 3. The upper bound follows from Lemma 4. \blacksquare

Computing $f(A^kM^{n-k})$ will be a bit more involved.

Proof of Theorem 3. The lower bounds follow from Lemmas 2 and 3. For the upper bound, we say that the i th row of a grid is a *top row* if $i \leq k$, a *middle row* if $k < i < n - k$, and a *bottom row* if $i \geq n - k$. We similarly define left, middle, and right columns. Let G be a grid. An observation that we use throughout this proof is that if $G(i, j) = G(i, j')$ with $j \leq k$ and $j' \geq n - k$, then the i th row cannot contain w (as we would need one of these entries to be A and the other to be M). Similarly if $G(i, j) = A$ with $k < j < n - k$, then the i th row cannot contain w . Analogous results hold with the roles of columns and rows swapped.

First consider the case that both diagonals of G contain w . By rotating G , we can assume $G(1, 1) = G(1, n) = A$ and that $G(n, 1) = G(n, n) = M$. Because $G(i, i) = G(i, n - i + 1) = A$ for $1 \leq i \leq k$, none of the k top rows can contain w , and similarly none of the k bottom rows can contain w . If no row contains w , then $f(w, G) \leq n + 2$. This equals $\max\{4k, 2(n - k) + 1\}$ only when $n = 2, k = 1$ or $n = 3, k = 1$, and for any other values it is strictly smaller. Thus we can assume that some number of the $n - 2k$ middle rows contain w , say with $G(i, j) = w_j$ for some $k < i < n - k$ and all j ; see Figure 7 for an example. This implies that none of the k left columns contain w , so in total counting rows, columns, and diagonals we have

$$f(w, G) \leq (n - 2k) + (n - k) + 2 \leq 2(n - k) + 1,$$

since $k \geq 1$.

Now assume exactly one diagonal of G contains w . By rotating G we can assume that this is the diagonal containing $(1, 1)$ and (n, n) and that $G(1, 1) = A$, $G(n, n) = M$. If some i th row of G with $i \geq n - k$ contains w , then we must have $G(i, j) = A$ for $j \leq k$ since $G(i, i) = M$ by assumption of the diagonal containing w . Because $G(i, j) = A$ for all $j \leq k$, we conclude that, as in Figure 8, none of the left columns can contain w if one of the bottom rows contains w . Similarly, either G contains none

A				A
	A		A	
A	A	M	M	M
	M		M	
M				M

Figure 7. Both diagonals filled in as well as one middle row. Now neither of the two left columns can contain AAMMM.

of the k right columns or none of the k top rows, so in total we have $f(w, G) \leq 2(n - k) + 1$.

A				
	A			
		M		
A	A	M	M	M
				M

Figure 8. One diagonal filled in as well as one bottom row. Neither of the left two columns can contain AAMMM.

Finally we assume no diagonal of G contains w . If none of the middle $n - 2k$ rows or columns contain w , then $f(w, G) \leq 2n - 2(n - 2k) = 4k$. If there exists $k < i, i' < n - k$ such that both the i th row and i' th column contain w , say with $G(i, j) = G(j, i') = w_j$ for all j , then, as in Figure 9, none of the k top rows or left columns can contain w , so $f(w, G) \leq 2n - 2k$. If none of the $n - 2k$ middle columns contain w and if there is some $k < i < n - k$ with $G(i, j) = w_j$ for all j , then none of the k left columns contain w and we have $f(w, G) \leq 2n - (n - 2k) - k = n + k$. Note that $n + k < 4k$ when $k > n/3$ and that $n + k < 2(n - k) + 1$ for $k \leq n/3$, so $f(w, G)$ is strictly smaller than the stated upper bound in this case.

		A		
		A		
A	A	M	M	M
		M		
		M		

Figure 9. Neither of the top two rows and neither of the left two columns can contain AAMMM.

■

By analyzing this argument carefully, it is not difficult to characterize all extremal examples for $A^k M^{n-k}$. In particular, Figure 2 contains all of the extremal constructions up to rotation and reflection of the grid. Thus, for the advertising job, one can freely choose whichever design one finds to be the most aesthetically pleasing. We personally recommend one of the middle grids, as these also contain a copy of the word MAA.

3. HIGHER DIMENSIONS.

Definitions and upper bounds. Given a set S and positive integer d , let S^d denote the set of sequences of length d where every entry is an element of S . Let $\vec{0}$ denote the sequence of d zeros, and if $x, y \in \mathbb{Z}^d$, define $(x + y) \in \mathbb{Z}^d$ by $(x + y)_i = x_i + y_i$.

Given $p \in [n]^d$ and $v \in \{-1, 0, +1\}^d$ with $v \neq \vec{0}$, we say that $\ell := \{p, p + v, \dots, p + (n - 1)v\}$ is a *line* if each of these elements is in $[n]^d$, and in this case we say that ℓ has initial point p and is in direction v . We first make a basic observation that we use throughout this section.

Lemma 7. *Let ℓ be a line with initial point p in direction v . Then $v_j = +1$ implies $p_j = 1$ and $v_j = -1$ implies $p_j = n$.*

Proof. Because every point of ℓ is in $[n]^d$, in particular we have $p_j, p_j + (n - 1)v_j \in [n]$. If $v_j = +1$ this is only possible if $p_j = 1$, and similarly $v_j = -1$ implies $p_j = n$. ■

The line ℓ with initial point p in direction v is the same as the line with initial point $p + (n - 1)v$ and direction $-v$. To have a unique identification¹ for each line, we say that $(p; v)$ is the *canonical pair* of the line ℓ if ℓ is the line with initial point p in direction v , and if the first nonzero coordinate of v is $+1$. We also say that p, v are the *canonical initial point* and the *canonical direction* of this line.

If ℓ is a line with canonical pair $(p; v)$, then we let $\ell_i = p + (i - 1)v$. We define the *weight* of a line ℓ to be the number of nonzero entries in its canonical direction. Note that there are no lines of weight 0 since we always require $v \neq \vec{0}$.

For example, when $d = 2$, the $2n$ rows and columns are lines of weight 1 with canonical pairs of the form $((i, 1); (0, 1))$ and $((1, i); (1, 0))$, and the two diagonals are lines of weight 2 with canonical pairs $((1, 1); (1, 1))$ and $((1, n), (1, -1))$. In particular, there are $2n + 2$ lines when $d = 2$. The following lemma generalizes these formulas to higher dimensions.

Lemma 8. *Let L_r denote the set of lines in $[n]^d$ of weight r and let $L = \bigcup_{r \geq 1} L_r$. Then for all $r \geq 1$,*

$$|L_r| = \binom{d}{r} 2^{r-1} n^{d-r},$$

$$|L| = \frac{1}{2}((n + 2)^d - n^d).$$

We note that the formula for L in Lemma 8 is known, see, e.g., [2, 9], but we include its short proof for completeness. Other results concerning combinatorial properties of lines in higher dimensions can be found in [2, 3, 7], for example.

¹This only makes sense for $n \geq 2$, as for $n = 1$ there is only one line, which is a point, so the direction is not well-defined.

Proof. Informally, we wish to define a map ϕ from lines ℓ to sequences of length d indicating how the coordinates of elements in ℓ change as we move along the canonical direction v . For example, if $d = 3$ and $\ell = \{(1, 1, n), (2, 1, n - 1), \dots, (n, 1, 1)\}$, we define $\phi(\ell) = (+, 1, -)$ to indicate that the first coordinate increases as we move along v , the second stays fixed at 1, and the third coordinate decreases.

To make this precise, let $S = \{1, 2, \dots, n, +, -\}^d$. We define a map $\phi : L \rightarrow S$ as follows. If ℓ has canonical pair $(p; v)$, we define $\phi(\ell)_i = p_i$ if $v_i = 0$, $\phi(\ell)_i = +$ if $v_i = +1$, and $\phi(\ell)_i = -$ if $v_i = -1$. It is not difficult to see that ϕ is an injective map, so it will be enough to determine the cardinality of its image when restricted to each set L_r .

Define $T_r \subset S$ to be the set of sequences with exactly r of its positions equal to \pm and such that the first \pm symbol to appear is a $+$. Because the canonical direction of a line in L_r has exactly r nonzero coordinates with its first nonzero coordinate equal to $+1$, we see that $\phi(L_r) \subset T_r$, and it is not difficult to see that every element of T_r is mapped to be a unique element of $\phi(L_r)$. Note that $|T_r| = \binom{d}{r} 2^{r-1} n^{d-r}$, as every element of T_r can be identified by choosing the r positions that are \pm , then choosing for all but the first of these positions which of the two symbols \pm to take on, and then choosing any element in $[n]$ for each of the remaining positions. This proves the first result.

The second result follows from the binomial theorem and summing the bound for each L_r . More directly, $\phi(L)$ is the set of sequences with at least one coordinate equal to \pm with the first of these being a $+$. There are $(n + 2)^d - n^d$ sequences with at least one \pm symbol, and exactly half of these sequences have the first such symbol as $+$. ■

We say that G is an (n, d) -grid (or simply a “grid” if n and d are understood) if it is a function from $[n]^d$ to a set of letters. Given a word w of length n and a grid G , we say that a line ℓ contains w if either $G(\ell_1) \cdots G(\ell_n) = w$ or $G(\ell_n) \cdots G(\ell_1) = w$. We let $f(w, G)$ denote the number of lines containing w in G , and we define

$$f(w, d) = \max_{G \text{ an } (n, d)\text{-grid}} f(w, G).$$

Observe that Lemma 8 immediately gives that $f(w, d) \leq \frac{1}{2}((n + 2)^d - n^d)$ for all w and d . We can improve upon this bound for nonsymmetric words with a generalization of Lemma 4.

Lemma 9. *Let w be a word with $w_i \neq w_{n-i+1}$ for some i . Then*

$$f(w, d) \leq \frac{1}{4}((n + 2)^d - (n - 2)^d).$$

Proof. Without loss of generality we assume that $i < n - i + 1$. The proof will proceed in four steps:

- 1: We define sets Q_p for certain $p \in [n]^d$. These Q_p correspond to corners of sub-hypercubes in $[n]^d$, and the p are the “upper-left” corners of these sub-hypercubes. The Q_p sets will partition $[n]^d$.
- 2: We classify every line in $[n]^d$ according to which Q_p it intersects. In particular, we show that for each line there is exactly one p such that the line intersects Q_p twice.
- 3: We determine an upper bound for the number of lines that can intersect each Q_p in two places and that can contain w .
- 4: We sum over the upper bounds in Step 3.

Step 1: Let P be the set of points $p \in [n]^d$ that do not have any position equal to $n - i + 1$. Let $I(p) = \{j : p_j = i\}$, and for each $p \in P$, define Q_p to be the set of points $p' \in [n]^d$ with $p'_j = p_j$ if $j \notin I(p)$ and $p'_j \in \{i, n - i + 1\}$ otherwise. In other words, Q_p is the set of corners of an $|I(p)|$ -dimensional sub-hypercube that has p as one of its corners.

Note that each $p' \in [n]^d$ is in exactly one Q_p set, namely for the p that has $p_j = i$ whenever $p'_j = n - i + 1$ and that has $p_j = p'_j$ otherwise. Also observe that $|Q_p| = 2^{|I(p)|}$, since each $p' \in Q_p$ has $p'_j \in \{i, n - i + 1\}$ for each $j \in I(p)$ and this completely determines p' . In particular, $|Q_p| = 1$ for all $p \in P$ such that p has no coordinate equal to i .

Step 2: Let L be the set of lines in $[n]^d$. Given an $\ell \in L$ with canonical pair $(p'; v)$, define

$$\phi(\ell) = \{p' + (i - 1)v, p' + (n - i)v\};$$

we call these two points the i -*dentifiers* of ℓ . (For example, the 1-*dentifiers* of ℓ are the endpoints of ℓ .) We claim that ϕ is a bijection from L to $\bigcup_{p \in P} \binom{Q_p}{2}$, where $\binom{Q_p}{2}$ denotes the set of two-element subsets of Q_p .

Let ℓ be a line with canonical pair $(p'; v)$, and define $p \in [n]^d$ by

$$p_j = \begin{cases} p'_j & v_j = 0 \text{ and } p'_j \neq n - i + 1, \\ i & \text{otherwise.} \end{cases}$$

We have $p \in P$ by construction.

If $p_j \neq i$, then by construction we have that $p_j = p'_j$ and $v_j = 0$. Therefore $p'_j + (i - 1)v_j = p_j$. On the other hand, suppose $p_j = i$. Then either $v_j \neq 0$, or $v_j = 0$ and $p'_j \in \{i, n - i + 1\}$. In the former case, Lemma 7 forces $p'_j \in \{1, n\}$ and therefore $p'_j + (i - 1)v_j \in \{i, n - i + 1\}$. In the latter case, it is clear that $p'_j + (i - 1)v_j = p'_j \in \{i, n - i + 1\}$. By definition, this means $p' + (i - 1)v \in Q_p$. Similarly, $p' + (n - i)v \in Q_p$. Thus $\phi(\ell) \in \binom{Q_p}{2}$ and ϕ maps to the desired codomain.

The map ϕ is injective since any two points of ℓ determine the line. It remains to prove that ϕ is surjective. Fix some $p \in P$ and $\{q, q'\} \in \binom{Q_p}{2}$. Observe that if $q_j \neq q'_j$ for some j , then $|q_j - q'_j| = n - 2i + 1$. Since $q \neq q'$, the sequence $v = \frac{1}{n - 2i + 1}(q - q')$ is a nonzero element of $\{-1, 0, 1\}^d$, and possibly by relabeling q, q' , we can assume that the first nonzero coordinate of v is $+1$. Define the point $p' := q - (n - i)v$. Note that when $v_j = +1$, we have $q_j = n - i + 1$, so $p'_j = 1$. Similarly, when $v_j = -1$ we have $q_j = i$ and $p'_j = n$. Thus $p' + (j - 1)v \in [n]^d$ for all $1 \leq j \leq n$ and we have that $(p'; v)$ is the canonical pair of some line ℓ with $\phi(\ell) = \{q, q'\}$, proving the surjectivity of the map.

Step 3: Let L_p be the set of lines whose i -*dentifiers* are in Q_p , i.e., L_p is the preimage of $\binom{Q_p}{2}$ under ϕ . Let $f(w, L_p)$ be the number of lines of L_p that contain w in G . Note that ϕ being a bijection implies that each line belongs to exactly one L_p set, so

$$f(w, G) = \sum_p f(w, L_p).$$

Given $p \in P$, let $Q_p^+ = \{q \in Q_p : G(q) = w_i\}$ and $Q_p^- = \{q \in Q_p : G(q) = w_{n-i+1}\}$. Because $w_i \neq w_{n-i+1}$ by assumption, for a line $\ell \in L_p$ to contain w ,

we must have one of its i -identifiers be in Q_p^+ and the other in Q_p^- . Because each $\{q, q'\} \in \binom{Q_p}{2}$ is the i -identifier of a unique line in L_p , we have

$$\begin{aligned} f(w, L_p) &\leq \left| \left\{ \{q, q'\} \in \binom{Q_p}{2} : q \in Q_p^+, q' \in Q_p^- \right\} \right| \\ &= |Q_p^+||Q_p^-| \leq \left(\frac{1}{2}|Q_p| \right)^2 = 4^{|I(p)|-1}, \end{aligned}$$

and in particular this integer is 0 when $|I(p)| = 0$.

Step 4: Let $P_r \subset P$ be the set of p with $|I(p)| = r$. Note $|P_r| = \binom{d}{r}(n-2)^{d-r}$, as a point in P_r is uniquely determined by first choosing which of its r coordinates are equal to i , and then each of the remaining coordinates can be any element in $[n] \setminus \{i, n-i+1\}$. We conclude that

$$f(w, G) = \sum_{r=0}^d \sum_{p \in P_r} f(w, L_p) \leq \sum_{r=1}^d \binom{d}{r} (n-2)^{d-r} \cdot 4^{r-1} = \frac{1}{4}((n+2)^d - (n-2)^d),$$

where the last step used the binomial theorem. ■

Lower bounds. We recall a standard combinatorial lemma.

Lemma 10. For d a positive integer and $x, y \in \mathbb{R}$, we have

$$\sum_{r \text{ odd}} \binom{d}{r} x^r y^{d-r} = \frac{1}{2}((x+y)^d - (-x+y)^d).$$

Proof. Using the binomial theorem, we see that the right-hand side of this equation is equal to

$$\frac{1}{2} \sum_r \binom{d}{r} (x^r y^{d-r} - (-x)^r y^{d-r}) = \sum_{r \text{ odd}} \binom{d}{r} x^r y^{d-r}.$$

With this we can prove our exact result for anti-symmetric words in all dimensions. ■

Proof of Theorem 5. The upper bound follows from Lemma 9. For simplicity we denote the two letters of w by ± 1 and we assume $w_1 = +1$. Let I denote the set of i such that $w_i = +1$, so that $w_i = -1$ for all $i \notin I$. Given a point $p \in [n]^d$, define $\sigma_I(p) = |\{i : p_i \in I\}|$, and define the grid G by $G(p) = (-1)^{\sigma_I(p)}$.

We claim that every line of G of odd weight contains w . Indeed, let ℓ be a line of odd weight r with canonical pair $(p; v)$, and recall that $\ell_i := p + (i-1)v$. Let us first assume $\sigma_I(\ell_1)$ is even, so that $G(\ell_1) = +1 = w_1$. Inductively assume $G(\ell_i) = w_i$. Note that as we go from ℓ_i to ℓ_{i+1} , a total of r coordinates will change, each going from either i to $i+1$ or from $n-i+1$ to $n-i$, with all other coordinates remaining the same. If $w_{i+1} = w_i$, then we either have $i, i+1 \in I$ or $n-i+1, n-i \in I$, and in either case we have $\sigma_I(\ell_{i+1}) = \sigma_I(\ell_i)$ and $G(\ell_{i+1}) = w_i = w_{i+1}$, as desired. If instead $w_{i+1} \neq w_i$, then either $i, n-i+1 \in I$ or $i+1, n-i \in I$. Depending on the case, we have $\sigma_I(\ell_{i+1}) = \sigma_I(\ell_i) \pm r$. Thus $G(p) = w_i(-1)^r = -w_i = w_{i+1}$, where we used the assumption that r has odd weight. Essentially the same proof works

when $\sigma_I(\ell_1)$ is odd, but with our reading the words backwards instead of forwards. By Lemmas 8 and 10 we have

$$f(w) \geq f(w, G) \geq \frac{1}{2} \sum_{r \text{ odd}} \binom{d}{r} 2^r n^{d-r} = \frac{1}{4} ((n+2)^d - (n-2)^d).$$

■

The idea of the proof of Theorem 4 is as follows. Define

$$\pi_i(p) := |\{j : p_j = i\}|, \quad \tau_i(p) := \pi_i(p) + \pi_{n-i+1}(p).$$

By looking at the problem through a probabilistic lens, one can show that with high probability a random line of $[n]^d$ has weight roughly $\frac{2d}{n+2}$. Moreover, for any point p on this line, with high probability there is a unique $i \leq \lceil n/2 \rceil$ such that $\tau_i(p) \approx \frac{4d}{n+2}$ with $\pi_j(p) \approx \frac{d}{n+2}$ for all $j \neq i, n-i+1$. Motivated by this, if w is symmetric, we will define our grid G by setting $G(p) = w_i$ whenever there is a unique $i \leq \lceil n/2 \rceil$ with $\tau_i(p)$ large. This will give us almost every line for symmetric words.

To get the bound for nonsymmetric words, we further need to decide whether a point p with $\tau_i(p) \approx \frac{4d}{n+2}$ should be assigned to w_i or w_{n-i+1} . We will make this decision based on the parity of

$$\sigma(p) := \sum_{i \leq n/2} \pi_i(p).$$

Note that this is a special case of the function $\sigma_I(p)$ used in the proof of Theorem 5 with $I = \lceil n/2 \rceil$ (and in fact, any such σ_I can be used to give the desired result). The reason we make this definition is because of the following.

Lemma 11. *Let $(p; v)$ be the canonical pair of a line ℓ of odd weight. Then*

$$\begin{aligned} \sigma(p) &\equiv \sigma(p + (i-1)v) \pmod{2} & i \leq \lceil n/2 \rceil, \\ \sigma(p) &\not\equiv \sigma(p + (i-1)v) \pmod{2} & i > \lceil n/2 \rceil. \end{aligned}$$

Heuristically this says that we can use the parity of $\sigma(p)$ to determine whether we are on the “left” or “right” side of our desired line. We note that for n odd, this lemma says nothing about $i = (n+1)/2$. This is fine because we can arbitrarily assign this point to be on the left or right side of the line since $w_i = w_{n-i+1}$ in this case.

Proof. Let $\pi'_j(p) = |\{k : p_k = j, v_k = 0\}|$. Thus for all i , $\pi_j(p + (i-1)v)$ will always be at least $\pi'_j(p)$ (since these positions stay fixed for all i) plus any of the positions that are made equal to j due to moving along v . By Lemma 7, if $v_k = +1$ then $(p + (i-1)v)_k = i$, and if $v_k = -1$ then $(p + (i-1)v)_k = n-i+1$. Thus if r is the weight of ℓ and s is the number of positions k with $v_k = +1$, we have

$$\begin{aligned} \pi_j(p + (i-1)v) &= \pi'_j(p) \quad \text{if } j \neq i, n-i+1, \\ \pi_i(p + (i-1)v) &= \pi'_i(p) + s, \\ \pi_{n-i+1}(p + (i-1)v) &= \pi'_{n-i+1}(p) + r - s. \end{aligned}$$

In particular, every term of $\sigma(p + (i-1)v) = \sum_{j \leq n/2} \pi_j(p + (i-1)v)$ will equal $\pi'_j(p)$ except for $j = \min\{i, n-i+1\}$, at which point it is equal to either $\pi'_j(p) + s$

or $\pi'_j(p) + r - s$ depending on which value j takes. Thus for $i \leq \lfloor n/2 \rfloor$ (and in particular for $i = 1$), we have

$$\sigma(p + (i - 1)v) = s + \sum_{j \leq n/2} \pi'_j(p) = \sigma(p),$$

and this second equality shows that, for $i > \lceil n/2 \rceil$, we have

$$\sigma(p + (i - 1)v) = r - s + \sum_{j \leq n/2} \pi'_j(p) = r - 2s + \sigma(p).$$

Because r is odd, this last quantity is of different parity from $\sigma(p)$. ■

We now begin the setup for the formal proof of Theorem 4. Let w be a word of length $n \geq 3$ and define

$$c := \frac{3.9d}{n + 2}, \quad k := \frac{2.3d}{n + 2}.$$

We say that a point p is a *counter-point* if

$$c < \tau_1(p) < 1.1c \quad \text{and} \\ \frac{.99(d - \tau_1(p))}{n - 2} \leq \pi_i(p) \leq \frac{1.01(d - \tau_1(p))}{n - 2} \quad \text{for all } 1 < i < n.$$

We define the grid G_w on points p with $\sigma(p)$ odd as follows. $G_w(p) = w_1$ if p is a counter-point, $G_w(p) = w_i$ if $\tau_1(p) \leq c$ and if $\tau_i(p) \geq k$ for a unique $1 < i \leq n/2$, and $G_w(p)$ is assigned arbitrarily for all other points with $\sigma(p)$ odd. If $\sigma(p)$ is even we let $G_w(p) = w_n$ if p is a counter-point, $G_w(p) = w_{n-i+1}$ if $\tau_1(p) \leq c$ and $\tau_i(p) \geq k$ for a unique $1 < i \leq n/2$, and $G_w(p)$ is assigned arbitrarily for all other points.

Our first goal is to show that most lines whose endpoints are counter-points contain the word w . To do this, we use a version of the Chernoff bound which can be found in [1], for example.

Lemma 12 (Chernoff bound). *Let $\text{Bin}(n, p)$ denote a binomial random variable with n trials and probability p , i.e., $\Pr[\text{Bin}(n, p) = k] = \binom{n}{k} p^k (1 - p)^{n-k}$. Then for all $\lambda > 0$ we have*

$$\Pr[|\text{Bin}(n, p) - pn| > \lambda pn] \leq 2e^{-\lambda^2 pn/2}.$$

That is, with high probability binomial variables are close to their expectation.

Lemma 13. *Let w be a word of length $n \geq 3$ and p a counter-point with $\tau_1(p) = r$. Then the number of lines ℓ in G_w that have initial point p and that contain w is at least $(1 - 2e^{-c/10000})2^{r-1}$. If $w_i = w_{n-i+1}$ for all i , then the number of lines is at least $(1 - 2e^{-c/10000})2^r$.*

Note that in the statement of the lemma we do not require p to be a canonical initial point.

Proof. Let $S \subset [d]$ be a set of coordinates with $|S| = s \geq .49r$ such that $j \in S$ implies $p_j \in \{1, n\}$. Note that there are exactly $\binom{r}{s}$ ways to choose such a set S with $|S| = s$ by the hypothesis $\tau_1(p) = r$. Given such an S , we form a line ℓ_S with initial

point p and direction v defined by $v_j = 0$ if $j \notin S$, and otherwise $v_j = +1$ if $p_j = 1$ and $v_j = -1$ if $p_j = n$.

For any i and $j \neq 1, n, i$, we have

$$\begin{aligned} \tau_j(p + (i - 1)v) &= \tau_j(p) \leq 2.02 \frac{d - r}{n - 2} < 2.02 \frac{d - c}{n - 2} \\ &= \frac{2.02d}{n + 2} \cdot \frac{n - 1.9}{n - 2} \leq \frac{2.222d}{n + 2} < k, \end{aligned}$$

where the first equality used similar reasoning as in the proof of Lemma 11, the first two inequalities used the definition of counter-points and $j \neq 1, n$, and the penultimate inequality used the fact that $n \geq 3$. We also have for $i \neq 1, n$ that

$$\tau_i(p + (i - 1)v) = \tau_i(p) + s \geq 1.98 \frac{d - r}{n - 2} + .49r.$$

Because $c < r < 1.1c$, this implies

$$\begin{aligned} \tau_i(p + (i - 1)v) &> \frac{1.98(d - 1.1c) + 0.49c(n - 2)}{n - 2} \\ &= \frac{d}{n + 2} \frac{3.891n - 8.3562}{n - 2} \\ &\geq \frac{3.3168d}{n + 2} > k. \end{aligned}$$

Finally, because $r < 1.1c$, we have for $i \neq 1, n$,

$$\tau_1(p + (i - 1)v) = r - s \leq 0.51r < c.$$

Thus $G_w(p + (i - 1)v) \in \{w_i, w_{n-i+1}\}$ for all i (this holds for $i = 1, n$ because p is a counter-point), with the exact value depending on the parity of $\sigma(p + (i - 1)v)$. If $w_i = w_{n-i+1}$ for all i , then ℓ_S contains w . If w is not symmetric and ℓ_S is of odd weight, then by Lemma 11 we will either have $\sigma(p + (i - 1)v) = w_i$ for all i if $\sigma(p)$ is odd, or $\sigma(p + (i - 1)v) = w_{n-i+1}$ for all i otherwise. In either case ℓ_S contains the word, so we conclude that ℓ_S always contains w if it has odd weight.

It remains to count how many lines contain w in each of these cases. For the non-symmetric case, each set S of odd size $s \geq .49r$ gives a distinct line ℓ_S containing w . Using Lemmas 10 and 12, we find that the number of such lines is

$$\begin{aligned} \sum_{s \geq .49r, s \text{ odd}} \binom{r}{s} &= 2^{r-1} - \sum_{s < .49r, s \text{ odd}} \binom{r}{s} \\ &\geq 2^{r-1} - 2^r \sum_{s < .49r} \binom{r}{s} 2^{-r} \\ &= 2^{r-1} - 2^r \cdot \Pr[\text{Bin}(r, .5) < .49r] \\ &= 2^{r-1} - 2^{r-1} \cdot \Pr[|\text{Bin}(r, .5) - .5r| > .01r] \\ &\geq (1 - 2e^{-r/10000})2^{r-1} \geq (1 - 2e^{-c/10000})2^{r-1}. \end{aligned}$$

The analysis for the symmetric case is essentially identical, the only difference being that we sum over all $s \geq .49r$ regardless of its parity. ■

We next wish to show that there are many counter-points.

Lemma 14. *Let w be a word of length $n \geq 3$. For $c < r < 1.1c$, the number of counter-points with $\tau_1(p) = r$ is at least*

$$(1 - 2de^{-d/40000(n+2)}) \binom{d}{r} 2^r (n - 2)^{d-r}.$$

Proof. Uniformly at random choose a point with $\tau_1(p) = r$ by first uniformly choosing a set $J \subset [d]$ of size r and uniformly assigning $p_j \in \{1, n\}$ for all $j \in J$, and then uniformly choosing $p_j \in \{2, \dots, n - 1\}$ for all other j .

Note that for $j \neq 1, n$ the distribution of $\pi_j(p)$ is exactly $\text{Bin}(d - r, \frac{1}{n-2})$, and thus the probability that any $\pi_j(p)$ is outside the range $\frac{.99(d-r)}{n-2}$ and $\frac{1.01(d-r)}{n-2}$ is at most $2e^{-(d-r)/20000(n-2)}$ by Lemma 12. Because $r < 1.1c$ and $n \geq 3$,

$$\begin{aligned} \frac{d-r}{n-2} &> \frac{d - 1.1 \frac{3.9}{n+2} d}{n-2} = \frac{d(n+2) - 4.29d}{(n-2)(n+2)} = \frac{d(n-2.29)}{(n+2)(n-2)} \\ &\geq 0.71 \cdot \frac{d}{n+2} \geq 0.5 \cdot \frac{d}{n+2}. \end{aligned}$$

Therefore the probability that any $\pi_j(p)$ is outside the range $\frac{.99(d-r)}{n-2}$ and $\frac{1.01(d-r)}{n-2}$ is at most $2e^{-(d-r)/20000(n-2)} \leq 2e^{-d/40000(n+2)}$. By the union bound, the probability that none of the $\pi_j(p)$ are outside this range, and hence the probability that p is a counter-point, is at least $1 - 2de^{-d/40000(n+2)}$.

Because we chose p uniformly among all points with $\tau_1(p) = r$, and because the number of such points is exactly $\binom{d}{r} 2^r (n - 2)^{d-r}$, we conclude that

$$\frac{|\{p : \tau_1(p) = r, p \text{ counter-point}\}|}{\binom{d}{r} 2^r (n - 2)^{d-r}} \geq 1 - 2de^{-d/40000(n+2)},$$

which gives the desired result. ■

We are now ready to prove Theorem 4.

Proof of Theorem 4. For $n = 2$, one can either have a word of the form $w = AA$ (in which case the grid that always maps to A has every line containing w), or of the form $w = AM$ (in which case Theorem 5 applies and gives the desired bound). From now on we assume $n \geq 3$.

Consider G_w as defined above. By using Lemmas 13 and 14 (and noting that the former lemma double counts lines since we do not consider canonical initial points), we find

$$f(w, G) \geq \frac{1}{2} \sum_{c < r < 1.1c} (1 - 2de^{-d/40000(n+2)}) \binom{d}{r} 2^r (n - 2)^{d-r} (1 - 2e^{c/10000}) 2^{r-1}$$

$$= (1 - 2de^{-d/40000(n+2)})(1 - 2e^{c/10000})\frac{1}{4}(n+2)^d \sum_{c < r < 1.1c} \binom{d}{r} \left(\frac{4}{n+2}\right)^r \left(\frac{n-2}{n+2}\right)^{d-r}$$

Note that this sum is equal to

$$\Pr \left[c < \text{Bin} \left(d, \frac{4}{n+2} \right) < 1.1c \right] \geq \Pr \left[\frac{3.9d}{n+2} < \text{Bin} \left(d, \frac{4}{n+2} \right) < \frac{4.1d}{n+2} \right],$$

where we used $4.1 < 3.9 \cdot 1.1$. By Lemma 12, this is greater than $1 - 2e^{-d/800(n+2)}$. Hence

$$f(w, G) \geq (1 - 2de^{-d/40000(n+2)})(1 - 2e^{c/10000})(1 - 2e^{-d/800(n+2)})\frac{1}{4}(n+2)^d.$$

Using this lower bound (and the fact that $c \rightarrow \infty$ as $d \rightarrow \infty$), as well as Lemma 9, we conclude for nonsymmetric words that $f(w, G) \sim \frac{1}{4}(n+2)^d$. For symmetric words, we use the symmetric case of Lemma 13 and an analogous argument to conclude the desired result. ■

4. A FEW SHORT WORDS ON SHORT WORDS. In this section, we informally discuss results for words of length k in (n, d) -grids. We define ℓ to be a line of length k in the obvious way, and, similarly for an (n, d) -grid G , we define what it means for this line to contain a word w of length k in a manner analogous to what we did before. We let $f(w, n, d) = \max_G f(w, G)$. One can show that the total number of lines in $[n]^d$ of length k is exactly

$$\frac{1}{2}((3n - 2k + 2)^d - n^d).$$

The proof is essentially the same as that of Lemma 8. The main difference is that before if we had $\phi(\ell)_i = +$ we knew that $p_i = 1$, but now it could be any value that is at most $n - k + 1$. Thus, instead of writing $+$, we should use the symbol p_i^+ to specify this information, and similarly one should use p_i^- instead of $-$.

Note that if k is much smaller than n and d is fixed, then asymptotically this number is $\frac{3^d - 1}{2}n^d$. This can be seen more directly by choosing a random point p in $[n]^d$ and observing that with high probability, p is one of the two endpoints of a line in direction v for all of the $3^d - 1$ possible nonzero directions v .

For the rest of this section we focus primarily on the case $d = 2$, with many of these ideas carrying over to larger d . The best generic bounds we have in this case are the following.

Theorem 6. *If w is a word of length k , then*

$$f(w, n, 1) \cdot (3n - 4k) \leq f(w, n, 2) \leq f(w, n, 1) \cdot 2n + 4 \sum_{i=k}^n f(w, i, 1).$$

Note that if $f(w, n, 1) \sim \alpha n$ and k is much smaller than n , then asymptotically the lower bound is $3\alpha n^2$ and the upper bound is $4\alpha n^2$, so this essentially solves the problem within a factor of $4/3$ for all words w .

Proof. Let G be an $(n, 2)$ -grid. By definition, the number of copies of w that G contains in its first row is at most $f(w, n, 1)$, so in total the total number of copies it

contains in its rows is at most $f(w, n, 1) \cdot n$. The same result holds for the number of copies of w appearing in a column. Similarly there are (at most) four “maximal” diagonal lines in G of length i for any $i \geq k$, see Figure 10, and within each of these diagonals, G contains at most $f(w, i, 1)$ copies of w . As every line of length k appears in a row, column, or maximal diagonal of length at least k , we conclude the upper bound.

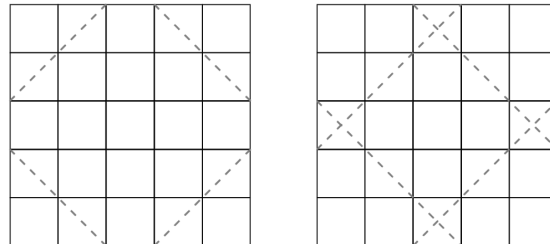


Figure 10. Dashed lines showing the four maximal diagonals of lengths 2 and 3 in a 5×5 grid.

For the lower bound, let G' be an $(n, 1)$ -grid that contains $f(w, n, 1)$ copies of w . Define $G(i, j) = G'(j)$. Observe then that the total number of copies of w appearing in rows is exactly $f(w, n, 1) \cdot n$. Moreover, if $v = (0, 1)$ and $p = (i, j)$ with $k \leq i \leq n - k$ is such that the line with initial point p and direction v contains w in G , then so do the lines with initial points p in directions $v' = (1, 1)$ and $v'' = (-1, 1)$. This gives an extra count of $f(w, n, 1) \cdot 2(n - 2k)$, as desired. ■

For example, if w' is the word of length $k \geq 2$ with k distinct letters, we claim that $f(w', n, 1) \approx n/(k - 1)$. For $w' = ABCD$, the corresponding lower bound construction looks like Figure 11:



Figure 11. A one-dimensional grid giving the maximum number of instances of the word ABCD.

In this grid, almost every copy of A and D is one of the two endpoints of two lines containing w' , so the total number of lines containing w' is roughly the number of times A and D appear in this grid. It is not hard to see that this happens for roughly one third of the positions in G , and a similar construction proves the lower bound of our claim for all k . For the upper bound, we have $f(w', k, 1) = 1$, and it is not difficult to see that for all n ,

$$f(w', n + k - 1, 1) \leq f(w', n, 1) + 1,$$

as the first n letters of an $(n + k - 1, 1)$ -grid contain at most $f(w', n, 1)$ copies of w' and at most one of the last $k - 1$ positions can be the endpoint of at most one line containing w' .

This claim together with Theorem 6 gives that $f(w', n, 2)$ is roughly between $\frac{3}{k-1}n^2$ and $\frac{4}{k-1}n^2$ when w' is a word on k distinct letters, and in general we do not know what the correct answer is for this very simple family of words, though we suspect that the lower bound is closer to the truth.

Conjecture 1. *If w' is the word of length $k \geq 2$ consisting of distinct letters, then*

$$f(w', n, 2) \sim \frac{3}{k-1}n^2.$$

In fact, it may be that $f(w, n, 2) \sim f(w, n, 1) \cdot 3n$ for any fixed word w .

5. APPENDIX. In order to get a better feel for the problem in higher dimensions, Figure 12 shows one of the three (up to rotation and reflection) optimal AMM grids for $d = 3$. More precisely it displays an “unfolded” two-dimensional version of the grid. To recover the three-dimensional grid, one can cut out the diagram and fold it into a cube along the gray lines (where the center of the cube is assumed to be M). There are 28 instances of AMM in this cube — can you see them all?

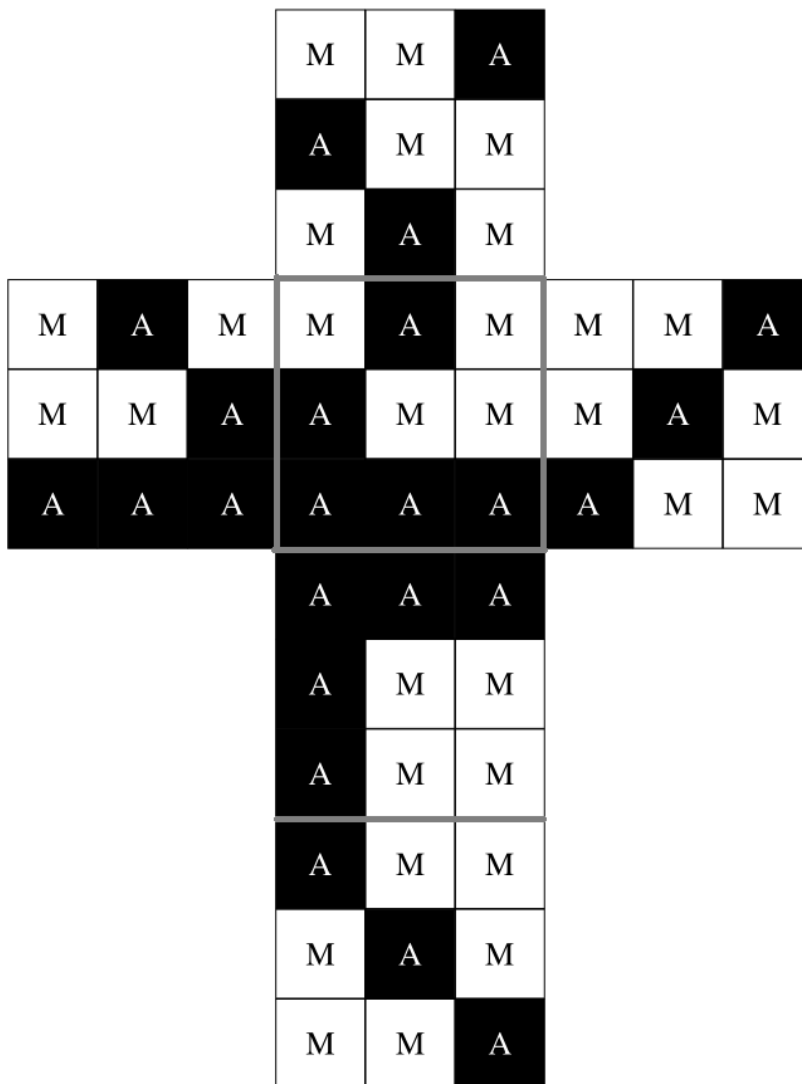


Figure 12. An unfolded $3 \times 3 \times 3$ grid giving the maximum number of instances of AMM.

ACKNOWLEDGMENTS. The authors are indebted to the two anonymous referees who greatly improved the quality of this paper. The authors wish to thank M. R. Thought for proposing this research problem and for fruitful discussions. The second author is supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1650112.

REFERENCES

1. Alon, N., Spencer, J. (2004). *The Probabilistic Method*, 2nd ed. New York, NY: Wiley–Interscience.
2. Beck, J. (2008). *Combinatorial Games: Tic-Tac-Toe Theory*. New York, NY: Cambridge Univ. Press.
3. Beck, J., Pegden, W., Vijay, S. (2009). The Hales–Jewett number is exponential: game-theoretic consequences. In: Chen, W. W. L., Gowers, W.T., Halberstam, H., Schmidt, W. M., Vaughan, R. C., eds. *Analytic Number Theory: Essays in Honour of Klaus Roth*, Vol. 14. Cambridge Univ. Press, pp. 22–37.
4. Dénes, J., Keedwell, D. (2015). *Latin Squares and Their Applications*, 2nd ed. Amsterdam: Elsevier.
5. Dénes, J., Mullen, G. (1993). Enumeration formulas for Latin and frequency squares. *Discrete Math* 111(1–3): 157–163. doi.org/10.1016/0012-365X(93)90152-J
6. Erdős, P., Hickerson, D., Norton, D., Stein, S. (1988). Has every Latin square of order n a partial Latin transversal of size $n - 1$? *Amer. Math. Monthly*. 95(5): 428–430. doi.org/10.1080/00029890.1988.11972024
7. Felzenbaum, A., Holzman, R., Kleitman, D. J. (1993). Packing lines in a hypercube. *Discret. Math.* 117(1–3): 107–112. doi.org/10.1016/0012-365X(93)90327-P
8. McKay, B., Wanless, I. (2008). A census of small Latin hypercubes. *SIAM J. Discret. Math.* 22(2): 719–736. doi.org/10.1137/070693874
9. OEIS Foundation Inc. (2020). The On-Line Encyclopedia of Integer Sequences. oeis.org/A005059
10. Ullrich, P (1999). An Eulerian square before Euler and an experimental design before R.A. Fisher: On the early history of Latin squares. *Chance*. 12(1): 22–26. doi.org/10.1080/09332480.1999.10542137
11. Zaikin, O., Kochemazov, S. (2015). The search for systems of diagonal Latin squares using the SAT@home project. *Int. J. Intell. Inf. Technol.* 3(11): 4–9.

GREGORY PATCHELL is a PhD student at UC San Diego studying operator algebras under Adrian Ioana. He is easily distracted by problems that are simple to state but hard to understand. Outside of math, his interests include drinking IPAs and avoiding barbers.

Department of Mathematics, UC San Diego, La Jolla CA 92093
gpatchel@ucsd.edu

SAM SPIRO is a PhD student at UC San Diego studying combinatorics with Jacques Verstraete. In addition to looking at words in grids, he also enjoys looking at words in books and manga.

Department of Mathematics, UC San Diego, La Jolla CA 92093
sspiro@ucsd.edu