A Rectangular Refinement of Aggarwal's Convex Unit-Distance Bound

ANONYMOUS MANUSCRIPT · DRAFT FOR EXTERNAL REVIEW
VERSION 0.1.1 · 21 JULY 2026

Anonymous manuscript — draft for external review

Version 0.1.1 — 21 July 2026

Review status. This manuscript contains a complete candidate derivation and reproducible finite checks. It has not been peer reviewed. The novelty claim is intentionally limited: no accepted or accessible prior statement of the explicit refinement below was found in the literature examined to date. A specialist in extremal 0–1 matrices or discrete geometry should verify both validity and prior art before public theorem language is used.

Abstract

Let $U_c(n)$ be the maximum number of unit-distance pairs among the vertices of a convex n-gon. Aggarwal proved the explicit upper bound

$$U_c(n)\le n\log_2 n+4n.$$

We give a rectangular refinement of the Tardos forbidden-submatrix estimate used in Aggarwal's proof. Retaining the two dimensions of the antipodal cut, rather than replacing both by their sum, yields the candidate bound

$$U_c(n)\le \left\lfloor 4n+\frac12\left[r\log_2(s-1)+s\log_2(r-1) +\log_2((r-1)!)+\log_2((s-1)!) \right] \right\rfloor,$$

where $r=\lfloor n/2\rfloor$ and $s=\lceil n/2\rceil$. Consequently,

$$U_c(n)\le n\log_2 n+ \left(3-\frac{\log_2 e}{2}\right)n+O(\log n),$$

whose linear coefficient is 2.278652479…, compared with 4 in the published explicit bound. The leading $n\log_2 n$ term is unchanged, so this does not resolve the Erdős–Moser conjecture that $U_c(n)=O(n)$. We also describe exhaustive small-matrix checks and arithmetic regression tests. These computations test finite

cases and implementation consistency; they are not a proof of the general statement.

1. The problem in plain language

Place n points at the corners of a convex polygon. Join two corners whenever they are exactly one unit apart. The question is: how many such joins can there be?

The best constructions currently listed have about $2n$ unit distances, while the best published upper bound has the larger order $n\log n$. Closing that gap is Erdős Problem #96. The present note does not close it. It trims the explicit linear term attached to the best published $n\log n$ upper bound.

The proof uses a bookkeeping device. Cut the polygon at two antipodal vertices—two vertices touched by parallel supporting lines—and lay the two boundary chains against the rows and columns of a grid. Put a 1 in a cell when its row vertex and column vertex are one unit apart. Geometry forbids a particular small arrangement of 1s. The geometric question can therefore be bounded by asking how many 1s a rectangular zero–one matrix may contain while avoiding that arrangement.

The improvement comes from preserving information that the published estimate discards. If the grid is a rows by b columns, replacing both dimensions by $a+b=n$ is like estimating the amount of paint needed for a rectangular wall by remembering only its perimeter. The perimeter is useful, but it forgets the wall's two side lengths. Here, keeping both side lengths allows the telescoping sums in Tardos's argument to be evaluated more sharply.

2. Background and notation

For a convex polygon P, write $U(P)$ for the number of unordered vertex pairs at Euclidean distance one, and define

```
U_c(n)=max{U(P): P is a convex n-gon}.
```

For zero–one matrices X and Q, say that X contains Q if some ordered submatrix of X can be changed into Q by turning selected 1s into 0s. Otherwise X avoids Q. Let $ex(R,C,Q)$ denote the maximum number of 1s in an $R\times C$ matrix avoiding Q.

Aggarwal associates an $a\times b$ cut matrix M_P to an antipodal cut of P, where $a+b=n$. Its 1 entries encode unit distances crossing the cut. Three published ingredients are used here.

1 The unit distances internal to the two boundary chains total at most $2n$ [1, Proposition 1].

2 The cut matrix avoids Aggarwal's pattern C', which is an amalgam of two smaller patterns A and B [1, Section 2.2]. Keszegh's amalgam inequality gives

$ex(a,b,C')\le ex(a,b,A)+ex(a,b,B)$ [2, Theorem 2.2; 1, Lemma 1].

1 Tardos bounds the relevant four-one pattern using two logarithmic weights that telescope across rows and down columns [3, Theorem 2.5 and the sharper observation following its proof].

All logarithms below are binary.

3. Main result

Theorem 1 (candidate rectangular refinement). Let $n\ge4$,

$$r=\lfloor n/2\rfloor,\qquad s=\lceil n/2\rceil.$$

Then

$$U_c(n)\le \lfloor 4n+\frac12[r\log_2(s-1)+s\log_2(r-1) +\log_2((r-1)!)+\log_2((s-1)!)] \rfloor.$$

Consequently,

$$U_c(n)\le n\log_2 n+ (3-(\log_2 e)/(2))n+O(\log n).$$

The proof is divided into four short lemmas. Lemma 1 preserves the rectangular dimensions in Tardos's argument. Lemma 2 obtains the transposed companion estimate. Lemma 3 inserts both estimates into Aggarwal's antipodal-cut proof. Lemma 4 shows that the resulting expression is largest when the two chains are as balanced as possible.

4. Rectangular Tardos estimate

The first pattern is

$$B= \begin{pmatrix} 1&0&1\\ 0&1&1 \end{pmatrix}.$$

Lemma 1. For $R\ge1$ and $C\ge2$,

$$ex(R,C,B) \le 2R+\frac12[R\log_2(C-1)+\log_2((C-1)!)].$$

Intuition. Ignore the first two 1s in every row. Each remaining 1 receives two charges. The first charge records how far the row has progressed from its first 1; those charges collapse, or telescope, to one endpoint ratio per row. The second charge records the complementary gap. Pattern avoidance forces those gaps to nest down each column, so those charges also telescope. Every retained 1 receives total charge at least two. Counting the available charge bounds the number of retained 1s.

Proof. Tardos's displayed pattern in the proof of Theorem 2.5 is a vertical reflection of B. Reversing the order of the rows preserves the rectangular extremal function, so it suffices to follow his proof for the reflected pattern.

Let $M=(m_{ij})$ be an $R\times C$ avoiding matrix. In row i, let f(i) be the column of the first 1. For a 1 in column j, let p(i,j) be the column of its preceding 1. Delete the first two 1s in every row, at a cost of at most 2R, and call the set of remaining 1-positions S. Thus

$$|S|\ge w(M)-2R,$$

where w(M) is the number of 1s in M.

For $(i,j)\in S$, define

$$w_1(i,j)=\log_2 (j-f(i))/(p(i,j)-f(i))$$

and

$$w_2(i,j)=\log_2 (j-f(i))/(j-p(i,j)).$$

Fix a row i. The w_1-terms telescope. If j_0 and j_1 are the last and second 1-positions in the row, then

$$\begin{aligned}\Sigma_j\ w_1(i,j)\\ =\log_2 (j_0-f(i))/(j_1-f(i))\\ \le \log_2 (C-1).\end{aligned}$$

Therefore

$$\Sigma_{(i,j)\in S}w_1(i,j)\le R\log_2 (C-1).$$

Fix a column j. For consecutive contributing rows $i<i'$, avoidance of the reflected pattern gives the nesting inequality

$$p(i,j)\le f(i').$$

The w_2-ratios therefore telescope down the column, exactly as in Tardos's argument, and give

$$\Sigma_i\ w_2(i,j)\le \log_2 (j-1).$$

Summing by columns yields

$$\begin{aligned}\Sigma_{(i,j)\in S}w_2(i,j)\\ \le \Sigma_{j=2}^{C}\log_2 (j-1)\\ =\log_2 ((C-1)!).\end{aligned}$$

Finally, put

$$t=(j-p(i,j))/(j-f(i))\in(0,1).$$

Then

$$w_1(i,j)+w_2(i,j)=-\log_2 (t-t^2)\ge 2,$$

because $t(1-t)\le 1/4$. Hence

$$2|S|\le R\log_2 (C-1)+\log_2 ((C-1)!).$$

Restoring the at most $2R$ deleted entries proves the lemma. □

5. Companion pattern

Aggarwal's second component pattern is

$$A=\begin{pmatrix}1&1\\ 1&0\\ 0&1\end{pmatrix}.$$

Lemma 2. For $a,b\ge 2$,

$$ex(a,b,A)\le 2b+\frac12[b\log_2 (a-1)+\log_2 ((a-1)!)].$$

Intuition. Pattern A is pattern B viewed after swapping rows with columns and reversing both orders. These operations change the orientation of the bookkeeping grid, not how many 1s it can contain.

Proof. Transposing B and reversing both row and column orders produces A. Applying the same operations to both host and pattern preserves ordered containment. Thus

```
ex(a,b,A)=ex(b,a,B),
```

and the claim follows from Lemma 1. □

6. Return to the polygon

Lemma 3. Let an antipodal cut of a convex n-gon have chain sizes $a,b\geq 2$, with $a+b=n$. Then

```
U(P)≤ 4n+\frac12[
alog₂(b-1)+blog₂(a-1)
+log₂((a-1)!)+log₂((b-1)!)
].
```

Intuition. The cut separates unit distances into three boxes: those entirely on the first chain, those entirely on the second chain, and those crossing between chains. Aggarwal bounds the first two boxes by $2n$ altogether. The two rectangular estimates bound the crossing box.

Proof. Aggarwal's forbidden C' pattern and Keszegh's amalgam inequality give

```
U(M_P)≤ ex(a,b,A)+ex(a,b,B).
```

Applying Lemmas 1 and 2,

```
U(M_P)≤ 2n+\frac12[
alog₂(b-1)+blog₂(a-1)
+log₂((a-1)!)+log₂((b-1)!)
].
```

Aggarwal's Proposition 1 adds at most $2n$ unit distances internal to the two chains, proving the displayed bound. If one chain has size one, the crossing contribution is at most $n-1$; the resulting elementary bound is smaller than the right-hand side of Theorem 1 for $n\geq 4$. □

7. Why the balanced cut is worst

Lemma 4. Subject to $a+b=n$ and $a,b\geq 2$, the variable part of the bound in Lemma 3 is maximized when $|a-b|\leq 1$.

Intuition. Imagine moving one vertex from the longer chain to the shorter chain. For the expression produced by the two telescoping estimates, that move always increases the bound until the two chain sizes are equal or differ by one. Thus the most demanding case is the balanced cut.

Proof. Exponentiate twice the variable part and define

```
F(a,b)=(b-1)^a(a-1)^b(a-1)!(b-1)!.
```

For $2\leq a\leq b-2$, direct cancellation gives

```
(F(a+1,b-1))/(F(a,b))
=(\frac a{a-1})^b
((b-2)/(b-1))^{a+1}.
```

For fixed a, each factor increases with b, so this ratio is at least its value at $b=a+2$:

$$\left(\frac{a}{a-1}\right)^{a+2}\left(\frac{a}{a+1}\right)^{a+1}=\frac{a}{a-1}\left(\frac{a^2}{a^2-1}\right)^{a+1}>1.$$

Moving one unit from the larger argument to the smaller therefore strictly increases F until the arguments differ by at most one. Since logarithm is increasing, the same is true of the variable part in Lemma 3. □

Proof of Theorem 1. Apply Lemma 3 and maximize with Lemma 4, taking $a=r$ and $b=s$. The number of unit distances is an integer, so the floor may be applied. Finally, Stirling's estimate

$$\log_2(m!)=m\log_2 m-m\log_2 e+O(\log m)$$

in the balanced expression gives

$$U_c(n)\le n\log_2 n+\left(3-\frac{\log_2 e}{2}\right)n+O(\log n).$$

□

8. What improved—and what did not

The published explicit upper bound is

$$n\log_2 n+4n.$$

The candidate refinement is

$$n\log_2 n+2.278652479\ldots n+O(\log n).$$

Thus the improvement is in the linear term:

$$\left(1+\frac{\log_2 e}{2}\right)n-O(\log n)=1.721347520\ldots n-O(\log n).$$

The leading $n\log n$ term remains. Therefore this result would sharpen the best currently listed published upper bound, but it would not prove the conjectured linear bound and would not solve Erdős Problem #96.

The structural lesson is also narrow. The gain comes from retaining heterogeneous dimensions through an existing proof: the row budget and the column budget are kept separate until the final optimization. This does not by itself provide a route past the logarithmic barrier.

9. Reproducible finite checks

The accompanying reference implementation evaluates the exact expression using $\log_2(m!)=\sum_{k=2}^{m}\log_2 k$. Its tests perform four checks.

1. Small forbidden matrices. Every zero–one matrix of sizes 2×3, 3×3, 3×4, and 4×3 is enumerated. Among matrices avoiding B, the observed maximum weight does not exceed Lemma 1's bound.
2. Balancing. For every $4\le n\le100$, every admissible split $a+(n-a)=n$ is evaluated and compared with the balanced split.

3. Numerical improvement. At $n=100,1000,10000$, the exact integer candidate bound is below $\lfloor n\log_2 n+4n\rfloor$.

4. Coefficient. The implementation checks

$$3-\frac{\log_2 e}{2}=2.2786524795555183\ldots.$$

These checks are useful because they can expose a mistranscribed pattern, a reversed inequality, a balancing error, or faulty arithmetic. They cannot establish the theorem for all dimensions. The general argument rests on the proofs above.

10. Novelty and review boundary

As of 21 July 2026, the Erdős Problems database lists the problem as open and records Aggarwal's $n\log_2 n+4n$ estimate as the best known upper bound [4]. The literature sweep conducted for this draft checked the exact coefficient, its decimal form, explicit improvements to Aggarwal's bound, rectangular refinements of the relevant Tardos estimate, and later work citing Aggarwal in related areas. No accepted or accessible source stating this explicit refinement was found.

That is evidence of possible novelty, not proof of novelty. The rectangular estimate may be unpublished folklore, may appear under different notation, or may be subsumed by a source not yet located. Before submission, two independent questions must be answered by a domain specialist:

1. Does every step of the rectangular telescoping, symmetry, amalgam application, and balancing argument hold as stated?
2. Is the explicit refinement already known in the literature or as folklore?

Until then, the correct description is: a candidate new refinement of the best currently listed published upper bound, pending expert review.

11. References

[1] A. Aggarwal, “On Unit Distances in a Convex Polygon,” Discrete Mathematics 338 (2015), 88–92. arXiv:1009.2216. https://arxiv.org/abs/1009.2216

[2] B. Keszegh, “On Linear Forbidden Submatrices,” Journal of Combinatorial Theory, Series A 116 (2009), 232–241. https://doi.org/10.1016/j.jcta.2008.05.006

[3] G. Tardos, “On 0–1 Matrices and Small Excluded Submatrices,” Journal of Combinatorial Theory, Series A 111 (2005), 266–288. https://doi.org/10.1016/j.jcta.2004.11.015

[4] T. F. Bloom, “Erdős Problem #96,” Erdős Problems, accessed 21 July 2026. https://www.erdosproblems.com/96

Appendix A. Minimal review checklist

- Reconstruct the column telescoping inequality in Lemma 1 directly from avoidance of the reflected pattern.
- Confirm that the row/column reversals in Lemma 2 map B to Aggarwal's A under the same containment convention.

- Confirm that Aggarwal's C' is the required amalgam and that the rectangular version of Keszegh's inequality applies.
- Check the one-vertex-chain boundary case.
- Recalculate the adjacent ratio in Lemma 4 and its monotonicity in b.
- Recompute the Stirling simplification, including the coefficient of n.
- Search explicitly for rectangular, non-square, or factorial-sum versions of Tardos's Q_1 estimate.

Appendix B. Disclosure for the review copy

Computational tools and language models assisted with literature retrieval, proof stress-testing, arithmetic checks, and drafting. The mathematical claim is presented only with an explicit derivation; automated outputs are not treated as mathematical authority. The manuscript remains anonymous for review, and identifying document metadata has been removed from the distributable file.