Finding patterns avoiding many monochromatic constellations

Steve Butler[*†] Kevin P. Costello[‡§] Ron Graham[¶]

Abstract

Given fixed $0 = q_0 < q_1 < q_2 < \cdots < q_k = 1$ a constellation in $[n]$ is a scaled translated realization of the q_i with all elements in $[n]$, i.e.,

$$p, p + q_1 d, p + q_2 d, \ldots, p + q_{k-1} d, p + d.$$

We consider the problem of minimizing the number of monochromatic constellations in a two coloring of $[n]$. We show how given a coloring based on a block pattern how to find the number of monochromatic solutions to a lower order term, and also how experimentally we might find an optimal block pattern. We also show for the case $k = 2$ that there is always a block pattern that beats random coloring.

1 Introduction

A constellation pattern is $\mathcal{Q} = [q_0 = 0, q_1, q_2, \ldots, q_{k-1}, q_k = 1]$ with q_i rational and $0 < q_1 < q_2 < \cdots < q_{k-1} < 1$. Given $\mathcal{Q}$ a constellation in $[n] = \{1, 2, \ldots, n\}$ is

$$p, p + q_1 d, p + q_2 d, \ldots, p + q_{k-1} d, p + d,$$

where each term is in $[n]$, or in other words a constellation in $[n]$ is a scaled and translated copy of the constellation pattern. We allow for d to be negative (i.e., the pattern to be reflected), so that we it does not matter whether we work with$[0, q_1, q_2, \ldots, q_{k-1}, 1]$ or $[0, 1 - q_{k-1}, 1 - q_{k-2}, \ldots, 1 - q_1, 1]$ (i.e., the mirrored version of the pattern).

The most well studied example of constellations are k-term arithmetic progressions, which correspond to the case $q_i = i/(k-1)$ for $i = 0, 1, \ldots, k$. Another example are solutions to equations of the form $ax + by = (a + b)z$ where $x, y, z \in [n]$, which corresponds to $[0, a/(a + b), 1]$.

For any constellation pattern $\mathcal{Q}$ we will let D be the smallest common denominator of the q_i. The number of constellations in $[n]$ is $n^2/D + O(n)$ if the pattern is not symmetric and $n^2/(2D) + O(n)$ if the pattern is symmetric. One way to see this is to pick two elements $p, q \in [n]$ (which can be done in n^2 ways), at which point p and q are the start and end of a constellation if and only if $D \,\big|\, (p - q)$, which happens with probability $1/D$, giving $n^2/D + O(n)$ constellations. When the pattern is symmetric we can interchange p and q so we divide by 2.

A natural question that arises is the following: Given a constellation $\mathcal{Q}$ and a fixed number r of colors, can we color $[n]$ in such a way as to avoid having a monochromatic constellation (i.e., one where all the $p + q_i d$ are colored with the same color)? The answer to this is a resounding no, in that not only must we have monochromatic constellations for n large, but a positive fraction of all constellations must be monochromatic.

Fact 1. *For any constellation pattern $\mathcal{Q}$ there is a constant $c(\mathcal{Q})$ so that for* any *r coloring of $[n]$ there are at least $c(\mathcal{Q})n^2$ monochromatic constellations.*

[*]UCLA, butler@math.ucla.edu
[†]Supported by an NSF Postdoctoral fellowship.
[‡]Georgia Institute of Technology, kcostell@math.gatech.edu
[§]Supported in part by NSF grant DMS-0456611
[¶]UCSD, graham@ucsd.edu

To see this, given $\mathcal{Q}$ we note that the constellation corresponding to the arithmetic progression of length $D+1$ contains all the q_i in $\mathcal{Q}$. In particular the constellation $\mathcal{Q}$ is contained in an arithmetic progression, so if there are monochromatic progressions there must also be monochromatic constellations. So it suffices to show that for any two coloring of $[n]$ that there must be at least $d(k)n^2$ monochromatic arithmetic progressions of length k where $d(k) > 0$ is some constant. This was proved for the case $k = 3$ by Frankl, Graham and Rödl [3]. The same proof works for arbitrary k, we include it here for completeness.

By a theorem of van der Waerden [9] there is number $W := W(r, k)$ so that any r coloring of $[W]$ must have a monochromatic k-term arithmetic progression. We first note that as done above that there are $n^2/\big(2(W+1)\big)$ different arithmetic progressions of length W inside $[n]$. Each one of these must contain a monochromatic progression in a r coloring of $[n]$. To correct for any overcounting we note that each monochromatic progression in $[n]$ will be counted at most $\binom{W}{2}$ times, since there are at most $\binom{W}{2}$ ways for us to put the progression into $[W]$. Therefore there are at least n^2/W^3 monochromatic progressions.

Since we must have some positive fraction of the constellations be monochromatic the next natural question is what is the smallest number of monochromatic constellations, i.e., the smallest coefficient γ so that there are $\gamma n^2 + o(n^2)$ monochromatic constellations for n arbitrarily large, and how do we achieve this lower bound.

One obvious candidate is to consider random colorings. Since a constellation with $k+1$ points will be monochromatic with probability $1/2^k$ then by coloring randomly we get a coefficient of $\gamma = 1/(2^k D)$ if the constellation pattern is not symmetric and $\gamma = 1/(2^{k-1}D)$ if the pattern is symmetric.

Parrilo, Robertson and Saracino [4] considered this problem for 3-term arithmetic progressions (which corresponds to the constellation $\mathcal{Q} = [0, 1/2, 1]$). They showed that by subdividing $[n]$ into 12 appropriately sized blocks that we can have $(117/2192)n^2 + O(n)$ monochromatic constellations. Note that $117/2192 \approx 0.05337591\ldots < 1/16 = 0.0625$, so their coloring has significantly fewer progressions than a typical random coloring (roughly 85.4% of what we would expect if we color randomly).

In this paper we show how one could find this subdivision for 3-term arithmetic progressions experimentally. We also generalize the approach for other constellations and find colorings which beat random for 4 and 5-term arithmetic progressions as well as other constellation patterns. We show that for the constellation $[0, q, 1]$ that there is a way of coloring of $[n]$ that beats random. We also relate some of our techniques to related problems not involving constellations and conclude with some open problems.

2 Finding a coefficient of a block coloring

Given a coloring of $[n]$ where there are large runs of a single color we naturally can group these runs into blocks. A block pattern $\mathcal{B} = \langle b_1, b_2, \ldots, b_m\rangle$ then represents the relative sizes of blocks to one another. Since we only care about the relative sizes of the blocks we can scale all numbers by any constant. As an example the block pattern found by Parrilo et al. is

$$\langle 28, 6, 28, 37, 59, 116, 116, 59, 37, 28, 6, 28\rangle.$$

Pictorially this is shown in Figure 1.

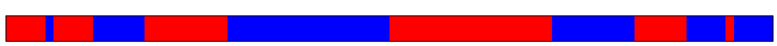


Figure 1: A good block coloring for avoiding 3-term arithmetic progressions.

Closely related to a block pattern is a subdivision pattern $\mathcal{X} = \langle\langle \beta_0, \beta_1, \ldots, \beta_m\rangle\rangle$ which gives the subdivision of the interval $[0, 1]$ according to the block pattern. It is easy to go back and forth between these two. Namely, given a block pattern then the subdivision pattern is found by letting

$$\beta_i = \frac{\sum_{j=1}^{i} b_j}{\sum_{j=1}^{m} b_j} \quad \text{for } i = 0, 1, \ldots, m,$$

while given a subdivision pattern to find the block pattern we let $b_i = \beta_i - \beta_{i-1}$ for $i = 1, 2, \dots, m$ and then, if desired, scale all the blocks by some constant.

Given a block pattern $\mathcal{B} = \langle b_i \rangle$ with corresponding subdivision pattern $\mathcal{X}$, the $\mathcal{B}$ coloring of $[n]$ is a two coloring found by coloring red all m with $\beta_{2i}n \le m \le \beta_{2i+1}n$, blue all m with $\beta_{2i-1}n \le m \le \beta_{2i}n$ and any blocks left over are colored arbitrarily.

Theorem 1. *Given a constellation pattern $\mathcal{Q} = [q_0, q_1, \dots, q_k]$ and a block pattern $\mathcal{B} = \langle b_1, \dots, b_m \rangle$. Then the number of monochromatic constellations of $\mathcal{Q}$ in a $\mathcal{B}$ coloring of $[n]$ is*

$$\begin{cases} \dfrac{\alpha}{2D} n^2 + O(n) & \text{if } \mathcal{Q} \text{ is symmetric,} \\ \dfrac{\alpha}{D} n^2 + O(n) & \text{if } \mathcal{Q} \text{ is not symmetric.} \end{cases}$$

Where

$$\alpha = \int_0^1 \int_0^1 \left(\prod_{i=0}^{k} \frac{1 + f\big(q_i x + (1-q_i)y\big)}{2} + \prod_{i=0}^{k} \frac{1 - f\big(q_i x + (1-q_i)y\big)}{2} \right) dy\, dx, \tag{1}$$

and

$$f(x) = \begin{cases} 1 & \text{for } \beta_{2i} \le x \le \beta_{2i+1}, \\ -1 & \text{for } \beta_{2i-1} \le x \le \beta_{2i}. \end{cases}$$

The function $f(x)$ is acting as an indicator function for whether we are in a red or a blue block. If we let $g(x, y)$ be the function inside the integral in (1), then $g(x, y)$ is also acting as an indicator function, but in this case it takes values 0 and 1, where $g(x, y) = 1$ if and only if x and y are (respectively) the start and end of a monochromatic constellation in $[0, 1]$. The basic idea is if we know where the monochromatic constellations of the block pattern in $[0, 1]$ are then we also know where the monochromatic constellations in $[n]$ are.

An important aspect about $g(x, y)$ is that it can only change value when (x, y) crosses a line of the form $q_i x + (1 - q_i)y = \beta_j$. In Figure 2 we have plotted the function for $Q = [0, 1/2, 1]$, where red indicates where $g(x, y) = 1$ and $f(x) = 1$, blue indicates $g(x, y) = 1$ and $f(x) = -1$, white indicates where the function is $g(x, y) = 0$. We have also drawn all the lines of the form $q_i x + (1 - q_i)y = \beta_j$. In particular, note that *every* region where $g(x, y) = 1$ will be a convex polygon.

Proof. Let $C(\mathcal{Q}, \mathcal{B}, n)$ be the number of monochromatic constellations of $\mathcal{Q}$ in a $\mathcal{B}$ coloring of $[n]$. We now approximate the integral for $g(x, y)$ in terms of $C(\mathcal{Q}, \mathcal{B}, n)$.

We make the following claim: p and q are the start and end of a monochromatic constellation in the $\mathcal{B}$ coloring of $[n]$ if $D \,\big|\, (p - q)$ and $g(x, y) = 1$ in a neighborhood around $(p/n, q/n)$. Similarly, p and q are not the start and end of a monochromatic constellation in the $\mathcal{B}$ coloring of $[n]$ if $B \not| \,(p - q)$ or $g(x, y) = 0$ in a neighborhood around $(p/n, q/n)$.

The divisibility condition follows from what was done in the introduction. The requirement that $g(x, y) = 1$ is to ensure that each $q_i p + (1 - q_i)q$ is in the same color class as p and q. The reason we insist it hold for a neighborhood is to avoid any ambiguity that occurs at the coloring on a border between blocks.

Subdivide $[0, 1] \times [0, 1]$ into squares of the form $[iD/n, (i+1)D/n] \times [jD/n, (j+1)D/n]$ for $0 \le i, j \le \lfloor n/D \rfloor$. The function $g(x, y)$ is not constant in a square and a small neighborhood of the square only if one of the lines $q_i x + (1 - q_i)y = \beta_j$ hits the square. Since there are $(m + 1)(k + 1)$ lines and each line can cross at most $2n/D$ squares it follows that there are at most $2(m + 1)(k + 1)n/D$ of the $(\lfloor n/D \rfloor)^2$ squares in our subdivision that are not constant in the square and its neighborhood.

Finally, by divisibility considerations each square contains D points which correspond to the start and end of constellations.

We now approximate α. We have that α is at least D^2/n^2 times the number of squares in the subdivision which are identically 1 in the square and a neighborhood. On the other hand counting monochromatic

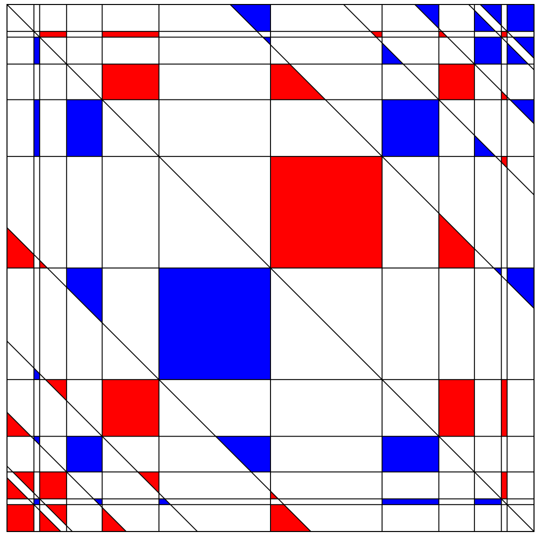


Figure 2: Indicator function for $\mathcal{Q} = [0, 1/2, 1]$ using the block pattern from Figure 1.

constellations we get D constellations for every such square and this misses at most $2(m+1)(k+1)n$ monochromatic constellations for squares we threw out which intersected a line. In particular we have that the number of squares is at least $(C(\mathcal{Q}, \mathcal{B}, n) - 2(m+1)(k+1)n)/D$. So we have

$$\alpha \geq \frac{D^2}{n^2} \frac{C(\mathcal{Q}, \mathcal{B}, n) - 2(m+1)(k+1)n}{D}$$

or rearranging,

$$C(\mathcal{Q}, \mathcal{B}, n) \leq \frac{\alpha}{D} n^2 + 2(m+1)(k+1)n.$$

A similar argument where we overcount monochromatic constellations and overestimate α gives

$$C(\mathcal{Q}, \mathcal{B}, n) \geq \frac{\alpha}{D} n^2 - 2(m+1)(k+1)n.$$

Combining the above two inequalities establishes the result for the non-symmetric case. For the symmetric case we divide by a factor of 2 because we restrict to the case when $p \leq q$. □

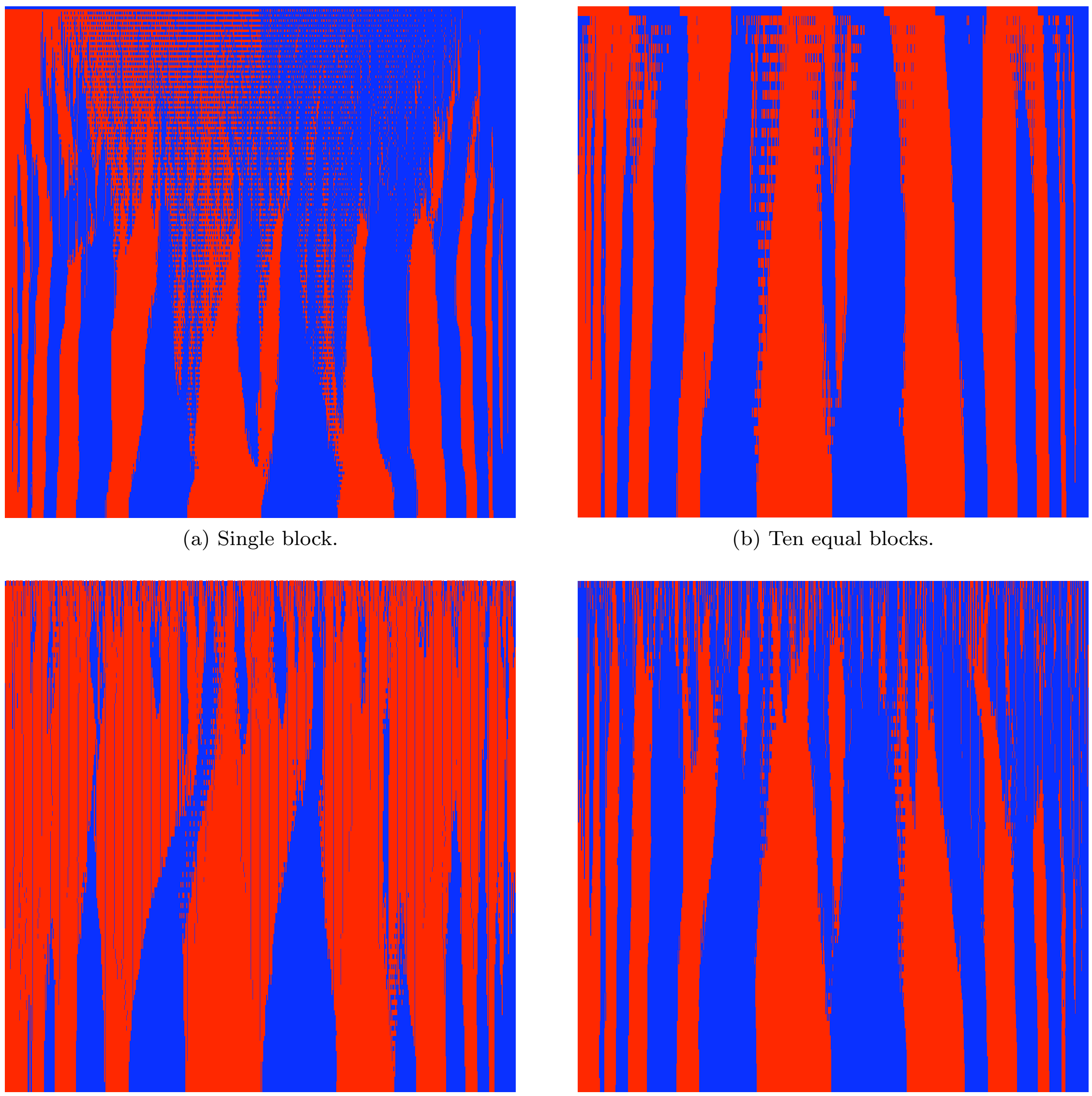
3 Perturbation to find good block patterns

Given a block pattern that colors $[n]$ we know now how to find the number of monochromatic constellations using the block coloring. To make use of this, we first need to find a good candidate block pattern. The goal of this section is to outline an approach for how such a pattern might be found. We will make use of the following observation twice.

Observation 1. *Given an optimal coloring for some fixed constellation pattern, then a small perturbation cannot decrease the number of monochromatic constellations.*

Let us first make use of the observation discretely. We fix n large, say 100000, and color $[n]$ arbitrarily. Now scan the elements of $[n]$, if we find an element where the switching the color decreases the number of monochromatic constellations then we switch and continue scanning. This process continues until we get to a coloring where changing the color on any single term will not decrease the number of monochromatic solutions, we will call such a coloring a locally optimal coloring on $[n]$. Note that one single element might change color multiple times in this process. However, the number of constellations strictly decreases on each pass, so the process will terminate in finite time.

When implementing this there are two major decisions, how to start the initial coloring; and how to scan for the next element to test for switching. In Figure 3 we show the evolution of a coloring on [1000] using several different starting colorings that converge to a locally optimal coloring for avoiding the constellation $[0, 1/3, 1]$ (this corresponds to avoiding monochromatic solutions to $x + 2y = 3z$). Our rule for scanning is to alternate between going left to right and right to left. We then output the current coloring when we hit the end of a row.



(a) Single block.

(b) Ten equal blocks.

(c) Random coloring.

(d) Random coloring.

Figure 3: Evolution of a locally minimal coloring for constellation $[0, 1/3, 1]$ with different starting colorings.

Note that in Figure 3 that starting with several different configurations they all converged to approximately the same block pattern that consists of 18 blocks. Intuitively, a block pattern that emerges by starting with $[n]$ and running this process should be an approximation to the optimal block pattern (if such a structure exists).

There are two problems. The first is that there is generally not a unique local optimal coloring, in other words there can be many patterns for which we cannot decrease the number of monochromatic constellations by changing the color of a single element. To deal with this we can run many iterations where after each iteration we flip some large fraction of the colors and run the process again. We then make some choice for which patterns are best, usually based on the ones having the fewest monochromatic constellations. While this does not guarantee that we find the best block structure it helps to rule out some possibilities.

The second problem is that the block pattern that we find is, at best, an approximate blowup of the optimal block pattern. For example if the best block structure has a block with small width, say $< 1/n$, then when we blow it up we might not catch the block in out pattern. To deal with this we generally choose n large (depending on the constellation). Another problem is that we do not have the precise relative sizes of the optimal block pattern. To deal with this we now perturb this near optimal block structure to settle into a locally optimal block structure, i.e., a block structure where an ϵ change in any of the β_i in $\mathcal{X}$ will increase the corresponding coefficient of the block coloring.

To do this we use the observation made earlier continuously. Namely, if our block structure is optimal then any small perturbation of the block sizes should increase the coefficient. This implies that if we look along the set of lines $q_i x + (1-q_i)y = \beta_j$ that a small change in β_j will add as much red (blue) as it will remove blue (red). If we are near the optimum then this allows us to set up a system of linear equations that the β_j in an optimal block structure must satisfy, i.e.,

$$\begin{pmatrix} \text{amount of change in red} \\ \text{under } \epsilon \text{ perturbation of } \beta_j \end{pmatrix} + \begin{pmatrix} \text{amount of change in blue} \\ \text{under } \epsilon \text{ perturbation of } \beta_j \end{pmatrix} = 0.$$

Where by the amount of change in red or blue we mean the change in area of all the polygons under a small ϵ perturbation of one of the β_j. For instance for the side of the polygon in Figure 4 we have

$$\Delta\text{Area} \approx \frac{\Delta x}{1-q_{j'}}\epsilon \approx \left(\frac{1}{q_{j'}-q_{i'}}\beta_i + \frac{1}{1-q_{j'}}\left(\frac{1-q_{k'}}{q_{j'}-q_{k'}} + \frac{1-q_{i'}}{q_{i'}-q_{j'}}\right)\beta_j + \frac{1}{q_{k'}-q_{j'}}\beta_k\right)\epsilon.$$

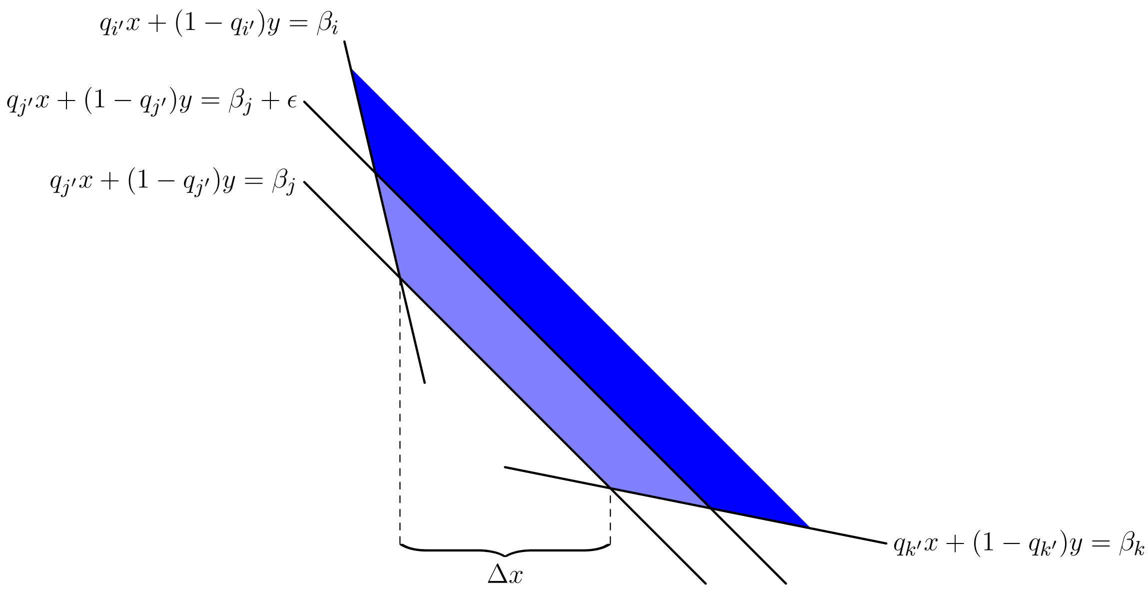


Figure 4: Change in area under an ϵ perturbation of β_j.

This must hold for $j = 1, 2, \ldots, k-1$, to this we also add $\beta_0 = 0$ and $\beta_k = 1$ to get a system of $k+1$ linear equations in $k+1$ unknowns. It is important that our block structure is near optimum since the set of linear equations is based on the polygons defining the regions where $g(x, y) = 1$ as used in Theorem 1. A different set of polygons in $g(x, y)$ lead to different equations that might produce an even worse coloring or even be undefined. (The important part of the polygons are the bounding lines of each polygon, so when we are near the optimal we do have the correct bounding lines and can perturb.)

(The process of setting up the linear equations can be completely automated and a `MAPLE` worksheet that implements this local perturbation is available at the first author's website.)

For example, for the constellation pattern $[0, 1/2, 1]$ doing a local perturbation on $[10000]$ we got an approximate block structure of

$$\langle 508, 109, 511, 674, 1076, 2116, 2117, 1077, 676, 512, 110, 514 \rangle.$$

Setting up the system of linear equations this corresponds to we get:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 6 & -12 & 10 & -6 & 6 & -6 & 2 & -2 & 2 & -2 & 2 & 0 & 0 \\ 2 & -10 & 16 & -10 & 6 & -6 & 2 & -2 & 2 & -2 & 4 & -2 & 0 \\ 4 & -6 & 10 & -14 & 10 & -6 & 2 & -2 & 2 & 0 & 2 & -2 & 0 \\ 2 & -6 & 6 & -10 & 16 & -10 & 2 & -2 & 4 & -2 & 2 & -2 & 0 \\ 4 & -6 & 6 & -6 & 10 & -14 & 6 & 0 & 2 & -2 & 2 & -2 & 0 \\ 2 & -2 & 2 & -2 & 2 & -6 & 8 & -6 & 2 & -2 & 2 & -2 & 2 \\ 0 & -2 & 2 & -2 & 2 & 0 & 6 & -14 & 10 & -6 & 6 & -6 & 4 \\ 0 & -2 & 2 & -2 & 4 & -2 & 2 & -10 & 16 & -10 & 6 & -6 & 2 \\ 0 & -2 & 2 & 0 & 2 & -2 & 2 & -6 & 10 & -14 & 10 & -6 & 4 \\ 0 & -2 & 4 & -2 & 2 & -2 & 2 & -6 & 6 & -10 & 16 & -10 & 2 \\ 0 & 0 & 2 & -2 & 2 & -2 & 2 & -6 & 6 & -6 & 10 & -12 & 6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \beta_6 \\ \beta_7 \\ \beta_8 \\ \beta_9 \\ \beta_{10} \\ \beta_{11} \\ \beta_{12} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Solving we get

$$\langle\!\langle \beta_i \rangle\!\rangle = \left\langle\!\!\left\langle 0, \frac{7}{137}, \frac{17}{274}, \frac{31}{274}, \frac{99}{548}, \frac{79}{274}, \frac{1}{2}, \frac{195}{274}, \frac{449}{548}, \frac{243}{274}, \frac{257}{274}, \frac{130}{137}, 1 \right\rangle\!\!\right\rangle,$$

which translating back into blocks gives us $\langle 28, 6, 28, 37, 59, 116, 116, 59, 37, 28, 6, 28 \rangle$. Now using this block structure we easily get our coefficient of $117/2192$.

This same technique can be used for any constellation (with the limitations mentioned above). For instance for the constellation $[0, 1/3, 1]$ doing local perturbation on $[25000]$ we get an approximate block structure

$$\langle 1101, 193, 577, 583, 989, 1434, 1115, 2833, 3680, 3681, 2830, 1113, 1434, 988, 582, 575, 194, 1098 \rangle.$$

Solving the system of linear equations we get a locally optimal block structure of

$$\begin{aligned} &\langle 1552213, 272415, 813251, 822338, 1394548, 2025068, 1572841, 3995910, 5196075, \\ &\ 5196075, 3995910, 1572841, 2025068, 1394548, 822338, 813251, 272415, 1552213 \rangle. \end{aligned}$$

giving a coefficient of $(16040191/211735908) \approx 0.075755\ldots < (1/12) \approx 0.083333\ldots$, showing that again in this case we can beat random.

For 4-term arithmetic progressions several different runs of $[100000]$ gave us an approximate block structure with 36 blocks. Perturbing, we found a pattern which has as a coefficient

$$\frac{17939629302218100912470205240133659380304674379 75}{1041774187682225982137537545158906769962544430213 44} \approx 0.0172202\ldots < \frac{1}{48} \approx 0.020833\ldots,$$

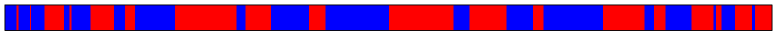


Figure 5: A good block coloring for avoiding 4-term arithmetic progressions.

again showing we can beat random (we give the corresponding block structure for this coefficient in the Appendix). The corresponding coloring is shown in Figure 5.

For 5-term arithmetic progressions, we found a block structure with 117 blocks which gave a coefficient of $0.005719619\ldots < (1/128) = 0.0078125$ showing yet again we beat random (the corresponding block structure is available at the first author's website). The corresponding coloring is shown in Figure 6.

Figure 6: A good block coloring for avoiding 5-term arithmetic progressions.

While we do not know if any of these block structures are optimal, we still note that the number of blocks seems to rise dramatically. We use 12, 36, and 117 blocks respectively for the colorings avoiding 3-, 4- and 5-term arithmetic progressions. In general we note that for a constellation with k points we need more than 2^{k-1} blocks in any block structure which beats random, so that the number of blocks needed grows exponentially with the number of points. To see this we note that the integral in Theorem 1 is at least as large as the area in the squares along the main diagonal. So if we have m blocks and we beat random then

$$\frac{1}{2^{k-1}} \geq \int_0^1 \int_0^1 g(x,y)\,dx\,dy \geq \sum_{i=1}^{m} \left(\beta_i - \beta_{i-1}\right)^2 \geq \frac{1}{m}.$$

In particular, this shows that the experimental approach runs into severe limitations as the number of points in the constellation gets large.

4 Solid blocks might not always be best

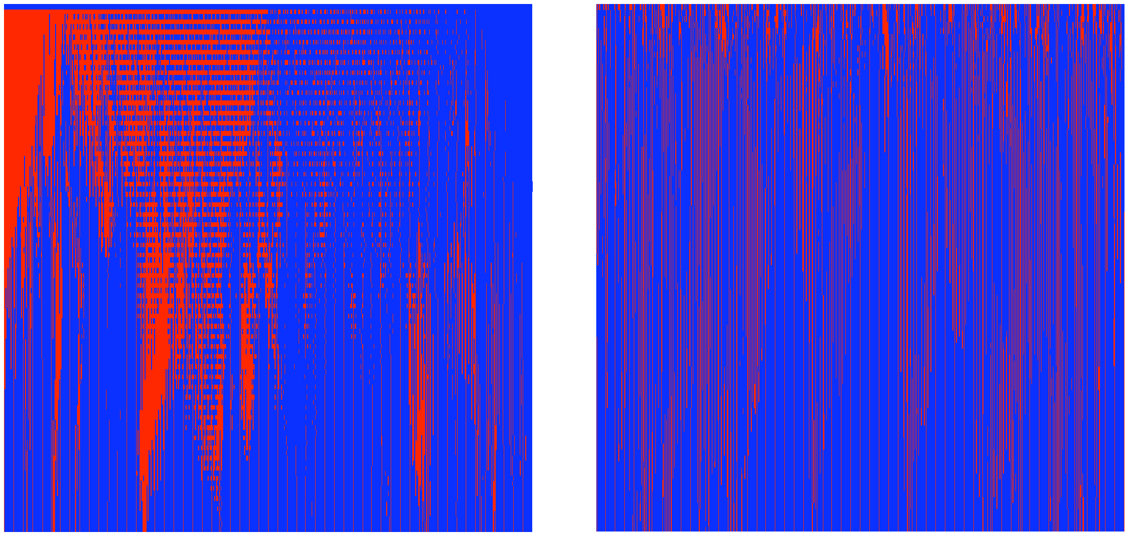
So far we have assumed that the optimal coloring of $[n]$ is done by blowing up large monochromatic blocks. But this might not always be the case. For example consider the constellation $[0, 2/5, 1]$ (which corresponds to avoiding monochromatic solutions to $2x + 3y = 5z$). In Figure 7 we show the evolution of a coloring on $[1000]$ to a locally optimal coloring for two starts (one monochromatic and one random).

The pattern that emerges in both of these cases (and many additional runs done for various block sizes, starts, and scanning rules) does not appear to be solid blocks but rather "alternating blocks", i.e., blocks which alternate red and blue in every entry, and between blocks there is an extra block, i.e.,

$$\begin{array}{c} \hphantom{\cdots RBRBRBRBR}\downarrow\hphantom{BRBRBRBRBR\cdots} \\ \cdots RBRBRBRBRRBRBRBRBRBR\cdots . \end{array}$$

This extra block has the property of shifting the modulus of the location of red and blue between two consecutive alternating blocks.

We can do the same process as before where we blow up a block pattern, but only make each block alternating and switch modulus between blocks. Also as before we can compute the coefficient that this corresponds to. The trick in doing this is to observe that a monochromatic constellation corresponds to a solution of $2x + 3y = 5z$ and if we look at the equations modulo 2 then we have $y \equiv z \pmod 2$. We can break our count into two situations, one where $x \equiv y \pmod 2$ and one where $x \not\equiv y \pmod 2$. The first case is counted as before, while the second case is counted by switching the color of the x term. This gives us a



(a) Single block. (b) Random coloring.

Figure 7: Evolution of a locally minimal coloring for constellation $[0, 2/5, 1]$ with different starting colorings.

coefficient κ where

$$\begin{aligned}
\kappa &= \frac{1}{10}\int_0^1\int_0^1 \left(\frac{(1+f(x))(1+f(y))(1+f\left(\frac{2x+3y}{5}\right))}{8} + \frac{(1-f(x))(1-f(y))(1-f\left(\frac{2x+3y}{5}\right))}{8}\right) dx\, dy \\
&\quad + \frac{1}{10}\int_0^1\int_0^1 \left(\frac{(1-f(x))(1+f(y))(1+f\left(\frac{2x+3y}{5}\right))}{8} + \frac{(1+f(x))(1-f(y))(1-f\left(\frac{2x+3y}{5}\right))}{8}\right) dx\, dy \\
&= \frac{1}{20}\int_0^1\int_0^1 \left(1 + f(y)f\left(\frac{2x+3y}{5}\right)\right) dx\, dy \\
&= \frac{1}{20} + \frac{1}{8}\int_0^1\int_{3y/5}^{(2+3y)/5} f(x)f(y)\, dx\, dy.
\end{aligned}$$

Where f is as in Theorem 1. The $1/20$ is fixed (and corresponds to the coefficient expected in a random coloring) so that our goal becomes to minimize the integral term which is an integral over a parallelogram. Since $f(x)f(y) = \pm 1$ then when we plot the function $f(x)f(y)$ we can mark where it is 1 by coloring blue and -1 by coloring red, see Figure 8 (it is important to note that red and blue signify something different than the red and blue given in Figure 2). Minimizing the integral then becomes equivalent to finding a pattern that maximizes the amount of red inside of the parallelogram.

Experimentally we find that an approximate (alternating) block pattern is

$$\langle 348, 113, 208, 325, 331, 731, 894, 731, 331, 325, 208, 113, 348\rangle.$$

As before we can locally optimize, which in this case means that for each β_j we must have that the amount of red immediately to the right of the line $x = \beta_j$ and above the line $y = \beta_j$ is equal to the amount of blue there (if this is not the case, we can increase the red by slightly increasing or decreasing β_j). As before this sets up a system of linear equations that can be solved to give a local optimum, doing so we get the following block pattern

$$\langle 9098298, 3018600, 5562432, 8660160, 8833560, 19511900, 23766825, \\ 19511900, 8833560, 8660160, 5562432, 3018600, 9098298\rangle$$

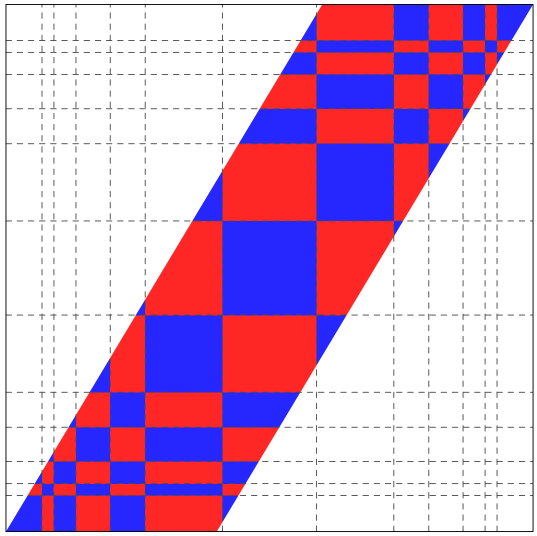


Figure 8: Maximizing the amount of red inside the parallelogram.

which gives a coefficient of $18447862/399410175 \approx 0.046187 \leq 1/20 = 0.05$ and again we have a coloring that beats random.

5 Beating the random coloring for constellations $[0, q, 1]$

In the previous sections we have seen colorings that beat random for the constellations $[0, 1/2, 1]$, $[0, 1/3, 1]$ and $[0, 2/5, 1]$. In this section we show that we can always beat random for any constellation of the form $[0, q, 1]$. We have already established the result for $[0, 1/2, 1]$ and by symmetry we only need to do the case $[0, q, 1]$ with $q < 1/2$, which is handled by the following.

Fact 2. *Let a, b be natural numbers with $2a < b$ and relatively prime and*

$$0 < \epsilon < 1 + \frac{a}{b} - \frac{a}{b}\left\lceil \frac{b}{a} \right\rceil.$$

Then for the constellation pattern $[0, a/b, 1]$ and the block pattern

$$\langle 1-\epsilon, 1+\epsilon, \underbrace{1, 1, \ldots, 1}_{2b-2 \text{ terms}} \rangle,$$

there are $\gamma n^2 + O(n)$ monochromatic constellations where

$$\gamma = \begin{cases} \dfrac{1}{4b} + \dfrac{(2a - a\lceil b/a \rceil)}{8ab^2(b-a)}\epsilon + O(\epsilon^2) & \text{if } \lceil b/a \rceil \text{ is odd,} \\ \dfrac{1}{4b} + \dfrac{(a - 2b + a\lceil b/a \rceil)}{8ab^2(b-a)}\epsilon + O(\epsilon^2) & \text{if } \lceil b/a \rceil \text{ is even.} \end{cases}$$

Since randomly we expect $(1/4b)n^2$ monochromatic constellations and in both cases above the coefficient for ϵ is negative, then for a small enough choice of ϵ the above subdivision pattern beats random. The key to this argument is that the block pattern with $\langle 1, 1, 1, \ldots, 1 \rangle$ with $2b$ blocks gives the coefficient $1/4b$ and so we only need to find a slight perturbation which would cause the coefficient to drop.

To see this note that by taking (1), expanding, and substituting this becomes

$$
\begin{aligned}
\frac{\alpha}{b} &= \frac{1}{4b}+\frac{1}{4b}\int_0^1\int_0^1 f(x)f(y)\,dx\,dy+\frac{1}{4b}\int_0^1\int_0^1 f(x)f\big(\frac{ax+(b-a)y}{b}\big)\,dx\,dy\\
&\qquad +\frac{1}{4b}\int_0^1\int_0^1 f(y)f\big(\frac{ax+(b-a)y}{b}\big)\,dx\,dy\\
&= \frac{1}{4b}+\frac{1}{4b}\int_0^1\int_0^1 f(u)f(v)\,du\,dv+\frac{1}{4(b-a)}\int_0^1\int_{av/b}^{(b+a(v-1))/b} f(u)f(v)\,du\,dv\\
&\qquad +\frac{1}{4a}\int_0^1\int_{(b-a)v/b}^{(bv-a(v-1))/b} f(u)f(v)\,du\,dv.
\end{aligned}
$$

Note the similarity to what we did in the previous section. In particular, calculating the coefficient reduces to calculating the difference between red and blue in the whole square and inside two parallelograms.

For the block pattern $\langle 1,1,1,\ldots,1\rangle$ with $2b$ blocks if we look at the inside integral of each term we see that the first one will look over the entire interval $[0,1]$, the second one will look over an interval of $[0,1]$ with width $(b-a)/b$ and the third one will look over an interval of $[0,1]$ with width a/b. Since the function f will change sign at regular steps of $1/2b$ it is easy to see that each of these inside integrals is 0. In particular for the block pattern all the integrals vanish and we are left with the constant term $1/4b$ which corresponds to random.

Now we simply perturb the pattern in the location of the first sign change of f, estimating the change of this perturbation to the integral reduces to estimating the difference between red and blue along the first line in the parallelograms, giving us the desired result.

The coloring we have produced in the above argument is almost certainly far from the best possible. To get a sense of how much better than random we can do we looked for optimal block colorings for $[0,q,1]$ for some simple q and plotted the ratio of the coefficient of this optimal coloring divided by the coefficient of the random coloring in Figure 9. The symmetry of the figure follows since we avoid $[0,q,1]$ when we avoid $[0,1-q,1]$ by reversing the coloring. We note that the lowest point is at the 3-term arithmetic progressions which corresponds to the ratio of $0.854\ldots$, also there seems to be a transition in behavior around $q=2/5,3/5$ which corresponds to the problem of avoiding monochromatic solutions of $2x+3y=5z$.

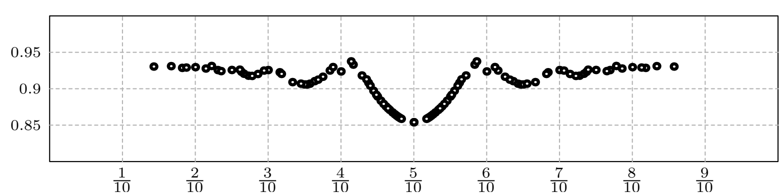
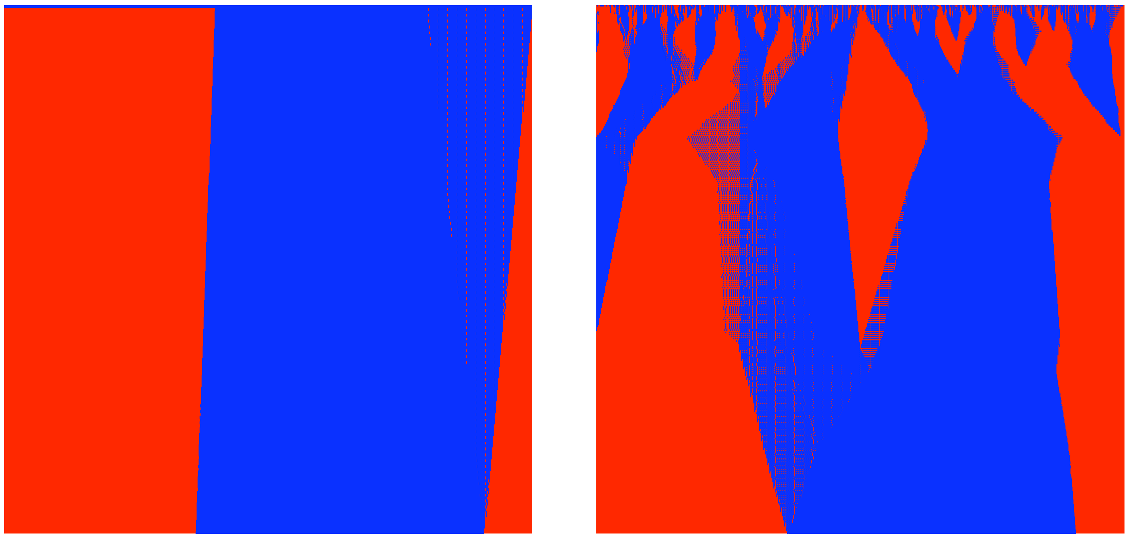


Figure 9: Good block colorings versus random for some $[0,q,1]$.

6 Non-constellation patterns

There are related questions of minimizing monochromatic solutions to equations where the solutions are not constellations. The most well-known example are Schur triples, which are solutions to $x+y=z$ (since solutions to this equation are not invariant under translation, they are not constellations). This is the currently only known situation where the minimal number of monochromatic solutions in a coloring of $[n]$ is known to be $(1/22)n^2+O(n)$ (see [2, 5, 6]). The lower bound achieving this is the block pattern $\langle 4,6,1\rangle$.

While Theorem 1 no longer applies in this situation (and so we cannot do local minimization of block structures), we can still experimentally find what should happen. For instance in Figure 10 we show the evolution of colorings on [1000] that avoid monochromatic solutions to $x+y=z$. In both of these runs (and many more) we see that the minimum has the form of a (medium block)–(large block)–(small block). Given this pattern it is not too hard to set up some variables for the three block sizes and to find the minimum which achieves the minimum number of monochromatic solutions.



(a) Single block. (b) Random coloring.

Figure 10: Evolution of a locally minimal coloring for Schur triples with different starting colorings.

We can also do the same process for other equations. For instance for $x + ky = z$, $k \geq 2$, experimentally we see that we get three blocks, again in the form medium–large–small. Suppose that we use the subdivision pattern $\langle\!\langle 0, \alpha, \beta, 1\rangle\!\rangle$. It is easy to show that the number of solutions in a monochromatic block $[p, q]$ (contained in $[n]$ with $(k + 1)p < q$) is

$$\frac{\big((q - kp) - p\big)^2}{2k} + O(n).$$

This gives $(\alpha^2/2k)n^2 + O(n)$ monochromatic solutions from the interval $[1, \alpha n]$ and $\big((\beta - (k+1)\alpha)^2/2k\big)n^2 + O(n)$ monochromatic solutions from the interval $[\alpha n, \beta n]$. The remaining solutions come from when x, z are in the third block and y in the first block, of which there are $\big((1 - \beta)^2/2k\big)n^2 + O(n)$ solutions. So altogether there are

$$\left(\frac{\alpha^2 + \big(\beta - (k+1)\alpha\big)^2 + (1 - \beta)^2}{2k}\right)n^2 + O(n)$$

monochromatic solutions. Optimizing our choice of α and β gives us a block pattern

$$\left\langle \frac{k+1}{k^2 + k + 3}, \frac{k^2 + k + 1}{k^2 + 2k + 3}, \frac{1}{k^2 + 2k + 3} \right\rangle,$$

which gives

$$\frac{1}{2k(k^2 + 2k + 3)}n^2 + O(n)$$

monochromatic solutions. (This same pattern was found independently by Thanatipanonda [7].)

For $ax + by = az$ with $a > b \geq 2$ and relatively prime, experimentally the optimal pattern appears to be to color $m \equiv 0 \pmod{a}$ red and the remaining terms blue, which gives a coefficient of $\big((2a - b)/(2a^4)\big)n^2 + O(n)$ monochromatic solutions. For $ax + by = az$ with $2 \leq a < b$ then experimentally the optimal pattern appears to be to color $m \equiv 0 \pmod{a}$ red for m small and the remainder blue. By optimizing like we have done above we have that we should color red for $m \equiv 0 \pmod{a}$ and

$$m < \frac{ab^2(a-1)}{b\big(b^2(a-1)+a\big)}n.$$

Doing this gives us

$$\frac{a-1}{2b\big(b^2(a-1)+a\big)}n^2 + O(n)$$

monochromatic solutions.

7 Concluding remarks

What we have done in the preceding sections is to give a systematic way to look for colorings having smaller than random number of monochromatic constellations and other patterns. This gives a way for giving upper bounds which we expect to be near-optimal for the minimum number of monochromatic constellations in such a coloring. However, there still remains the question of determining corresponding lower bounds. The only pattern for which the best known coloring matches (up to lower order terms) the best known lower bound are the Schur triples. (The lower bound given in the introduction is fairly weak and makes a poor candidate.)

For 3-term arithmetic progressions there is a lower bound of $(1675/32768)n^2 + o(n^2)$ given by Parrilo et al., which differ by about 5% from the previously mentioned upper bound. We believe that the correct value for 3-term progressions is the one given by the known locally-optimal coloring, i.e., $(117/2192)n^2 + O(n)$. Through several hundred runs in various sizes, starts and scanning rules the same block pattern came over and over again.

An interesting problem related to 3-term arithmetic progressions is the following: For a partition $0 = a_0 < a_1 < a_2 < \cdots < a_\ell = 1/2$ create a *checkerboard* pattern by taking a square of side lengths 1/2 and coloring the rectangle $[a_i, a_{i+1}] \times [a_j, a_{j+1}]$ blue if $i + j$ is even and red if $i + j$ is odd. What is the maximum amount of red that can be enclosed in the triangle with vertices at $(0,0)$, $(0,1/2)$, and $(1/2,1/4)$, and what partition (if any) produces this maximum? The best known pattern is found by scaling the block pattern $\langle 28, 6, 28, 37, 59, 116\rangle$ and is shown in Figure 11. (The connection is seen by looking at the integral for the case $q = 1/2$ in Section 5, and assuming the pattern is antisymmetric, i.e., $f(1-x) = -f(x)$.)

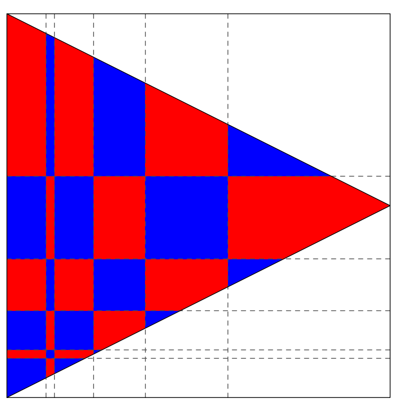


Figure 11: Best known pattern maximizing red inside of the triangle

Any improvement on this pattern would automatically produce a lower constant for the number of 3-term arithmetic progressions. By an exhaustive computer search there is no block pattern with 11 or fewer blocks that beats this pattern. This is additional evidence to support that the current pattern for 3-term arithmetic progressions is the best optimal.

One striking thing to notice is for $[0, 1/2, 1]$, $[0, 1/3, 1]$ and $[0, 2/5, 1]$ that all of the patterns we found are anti-symmetric, that is the color of i and $n + 1 - i$ are opposite. This same behavior occurs frequently for many (but not all) of the locally optimal block colorings that we found for constellations $[0, q, 1]$, it would be interesting to know if there is a reason that anti-symmetry is common.

In a related question, it would be interesting to know why $[0, 2/5, 1]$ goes into large alternating blocks. More generally, we might have a coloring where a block pattern emerges only when we look at what is happening modulo some appropriate p. Is there a way to predict beforehand, given a constellation pattern, whether the optimal coloring consists of solid blocks or some sort of alternating structure? Perhaps even more basic, is there a reason why we should expect block structures?

We have also seen that for the case $[0, q, 1]$ that we can always beat a random coloring. We conjecture that this holds in general.

Conjecture 1. *For any constellation pattern $\mathcal{Q}$ there is a coloring pattern of $[n]$ which has $\gamma n^2 + o(n^2)$ monochromatic constellations where γ is smaller than the coefficient for a random coloring.*

This is related to an idea in Ramsey theory where for some time it was thought that the best way to avoid monochromatic K_ts in a 2-coloring of K_n was to color randomly. Thomason [8] showed that this was not the case and produced colorings which beat random.

The key to proving the special case $[0, q, 1]$ was that we had a simple coloring that had the same number of monochromatic constellations as a random coloring, which we could then perturb. A first step in trying to prove the conjecture would be trying to find some "simple" block pattern that *matched* random, and then try and perturb it. This is not trivial, even for 4-term arithmetic progressions no simple pattern is known.

Another idea might be to try and bootstrap our way up. For instance one would expect that since every 4-term arithmetic progression has a 3-term arithmetic progression inside that by using the pattern for avoiding 3-term arithmetic progressions we would also avoid many 4-term arithmetic progressions. However, this is not the case, using the pattern for 3-term arithmetic progressions we do worse than random coloring for avoiding 4-term arithmetic progressions, and also vice-versa.

One might also consider this problem for two colorings of $\mathbb{Z}_n$. Cameron, Cillereulo and Serra [1] showed that for the constellations of the form $[0, q, 1]$ that the number of monochromatic constellations depends only on the amount of each color used, and not the distribution of the coloring in $\mathbb{Z}_n$. They also gave some lower bounds for the number of 4-term arithmetic progressions in colorings of $\mathbb{Z}_n$ which have been improved by Wolf [10].

Finally, by using roots of unity it is not hard to adapt Theorem 1 to the case when we have $r > 2$ colors. However, we have found enough beauty and mystery in the $r = 2$ case to keep us occupied for some time. We hope to see some of the problems addressed above in future works.

References

[1] Peter Cameron, Javier Cilleruelo and Oriol Serra, *On monochromatic solutions of equations in groups*, Rev. Mat. Iberoamericana **23** (2007), 385–395.

[2] Boris Datskovsky, *On the number of monochromatic Schur triples*, Advances in Applied Math. **31** (2001), 193–198.

[3] Peter Frankl, Ron Graham and Vojtěch Rödl, *Quantitative theorems for regular systems of equations*, J. Combin. Theory Ser. A **47** (1988), 246–261.

[4] Pablo Parrilo, Aaron Robertson and Dan Saracino, *On the asymptotic minimum number of monochromatic 3-term arithmetic progressions*, J. Combin. Theory Ser. A **115** (2008), 185–192.

[5] Aaron Robertson and Doron Zeilberger, *A 2-coloring of* $[1, N]$ *can have* $(1/22)N^2+O(N)$ *monochromatic Schur triples, but not less!*, Electron. J. Combin. **5** (1998), Research Paper 19, 4pp.

[6] Tomasz Schoen, *The number of monochromatic Schur triples*, Europ. J. Combinatorics **20** (1999), 855-866.

[7] Thotsaporn Thanatipanonda, *On the Monochromatic Schur Triples Type Problem*, Electron. J. Combin. **16** (2009), Research Paper 14, 11pp.

[8] Andrew Thomason, *A disproof of a conjecture of Erdős in Ramsey theory*, J. London Math. Soc. **39** (1989), 246–255.

[9] Bartel Leendert van der Waerden, *Beweis einer Baudetschen Vermutung*, Nieuw Arch. Wisk. **15** (1927), 212–216.

[10] Julia Wolf, *The minimal number of monochromatic* 4*-term progressions*, preprint.

Appendix

The locally optimal block pattern for 4-term arithmetic progressions which gives the coefficient of

$$\frac{1793962930221810091247020524013365938030467437975}{10417741876822259821375375451589067699625444302134 4}$$

is given using the following 36 blocks

$$\begin{aligned}
\langle & 566124189415440472939626834822903743300467940483,\\
& 115903533761943477398551818347715476722877927241,\\
& 568011813340950665677009286694526323061781532322,\\
& 472083073090028493914605548954507028673457587863,\\
& 174690683867336844297305424871758992360029965453,\\
& 98464537567500111285074159909918993309405119848,\\
& 737681146409933099806596775369238915383890216793,\\
& 881071132072892536672404740128385947619685842609,\\
& 387204684955306822603896642766296540832568256888,\\
& 340852889156784985628080980878507258595675472221,\\
& 1398355239284691808801098670395696996980804292522,\\
& 2015438904391090234472652819593929714355629836078,\\
& 354924006068259988552316716495705798216298952575,\\
& 917029329994691011286378833655488533756529343857,\\
& 1246774229265384930907724953794401373314144038191,\\
& 543203071437439856124749368271693956037186323582,\\
& 2179716742907087903057122171392866104311765026441,\\
& 2172387005301046067153961748343296914044366107569,\\
& 546203621232713973260465876924982631234637232779,\\
& 1296607046453245562932414262768745367411919249702,\\
& 848633230480614872785768439513578746631939174778,\\
& 332362434790023921274974476865878572983006589230,\\
& 2079963873190082657423397539748746717308015584742,\\
& 1352139932444260494597496699603210730948918467199,\\
& 339606780510267312616862401149984633870549046619,\\
& 373718051493152648659948917556716014372715915533,\\
& 786614601718483336599780069288734659594156639775,
\end{aligned}$$

$$
\begin{array}{l}
660138925526725209837882202781409057453701881412,\\
515057178882234589660036450164525746471722 14751,\\
208563593370975774208121482104389596165334050300,\\
660659764939424259451601477074423264167731648445,\\
458220356230536594713018106829384942175966684332,\\
251064024444857725679270083611802851025086 19312,\\
396852300398188165681981456085190506968094545183,\\
594925248577123140990661679710465570788206 29622,\\
398049321798723913182570904641775780071858799086\rangle
\end{array}
$$

The corresponding diagram for the function $g(x, y)$ from Theorem 1 is shown in Figure 12.

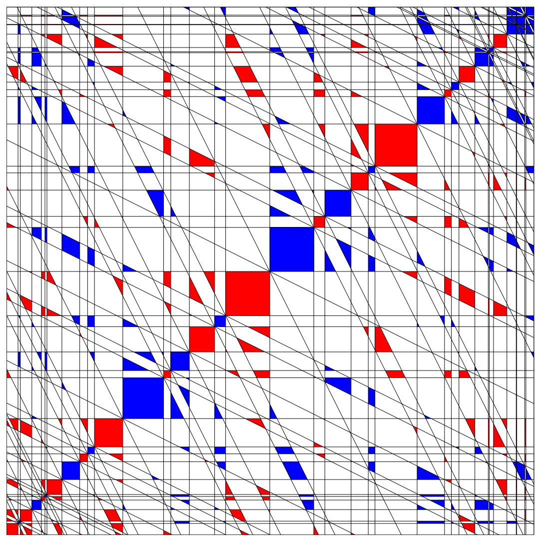


Figure 12: Indicator function for $\mathcal{Q} = [0, 1/3, 2/3, 1]$ using block pattern given in the Appendix.