# Property B: Random Greedy 2-coloring

Tibor Szabó

Extremal Combinatorics, FU Berlin, WiSe 2017–18

First let us recall the lower bound of Erdős on the number $m_B(k)$: Colour the vertices of $\mathcal{H}$ uniformly at random. If there are less than $2^{k-1}$ edges, then the expected number of monochromatic edges is less than 1 (by linearity of expectation). If there are more edges then we expect to see more monochromatic edges.
The first improvements of this lower bound were based on the so called alteration method: roughly speaking, if we are given a random colouring with few monochromatic edges, we can try to fix it by recolouring some vertices in the monochromatic edges. Note that one has to be careful: a potential fix could create further monochromatic edges. The idea is to do the fix randomly as well: when necessary, recolor vertices randomly. The analysis becomes quite delicate due to the dependece of the coloring and recoloring on each other. In an influential 1978 paper Beck managed to carry out an analysis and showed $m_B(k) \geq k^{\frac{1}{3}-o(1)}2^k$. In 2000 Radhakrishnan and Srinivasan improved this to $m_B(k) \geq k^{\frac{1}{2}-o(1)}2^k$. Although their proof is less complicated than Beck's, technically is still somewhat delicate.
Here first we discuss a bound weaker than Beck's, achieved more than three decades later, but by an incredibly simple approach.

**Theorem 0.1** (Pluhár, 2009)**.** $m_B(k) \geq ck^{\frac{1}{4}}2^k$

So what is the idea? Recall the *greedy coloring* that we used in Discrete Math I to produce a proper $(\Delta + 1)$-coloring for any graph $G$ of maximum degree $\Delta$. This was the procedure which passed through the vertices in some (arbitrary) order and colored each vertex with the "smallest available color". How does this procedure look like when we only have two colors? Well, we always color a vertex `blue` (the color 1), unless this would create a monochromatic edge, in which case we color it `red` (the color 2).
Note that this algorithm *never* creates a monochromatic `blue` edge, but it could create a monochromatic `red` edge. In fact there is a $k$-graph with only $k + 1$ edges, such that the greedy coloring might fail. (**HW**) This is despite that we already know that for every graph containing $k + 1$ edges, even $2^{k-1}$ edges *there is* a `red`/`blue`-coloring avoiding monochromatic edges. To avoid the failure or make sure that it does not likely happen, we will select the *order* in which we color the vertices not arbitrarily, but randomly.

*Proof.* Algorithm:
Input: a $k$-graph $\mathcal{H}$ on $n$-vertices with m edges.

Step 1: Choose an ordering of the vertices $v_1, ..., v_n$ uniformly at random;

Step 2: Consider the vertices in the order chosen:

(i) If the vertex is the last vertex in all `blue` edge, colour it `red`

(ii) Otherwise colour it `blue`.

Output: The colouring obtained.

Due to step 2(i) there will be no monochromatic `blue` edge, but there could be a monochromatic `red` edge $f$. Consider the first vertex in $f$. We colored this vertex `red` for a reason: it must be the last vertex in an otherwise all-`blue` edge $e$. So in particular, the Greedy Coloring Algorithm might fail coloring $f$ properly *only if*

> there is another edge $e$ that intersects $f$ in exactly one vertex which is in the middle ($k$th) position among the $2k-1$ elements of $e \cup f$.

Our analysis will focus on forbidding the occurrence of such pair of edges.

- We will bound the probability that the first vertex of one edge is the last vertex of another edge. (Without this, we cannot have an all `red` edge.)
- Events: for edges $e, f \in \mathcal{H}$, let $\mathcal{E}_{e,f}$ be the event that the last vertex of $e$ is the first vertex of $f$.
  (i) {algorithm fails} $\subseteq \bigcup_{e,f} \mathcal{E}_{e,f}$, hence $\mathbb{P}(\text{algorithm fails}) \leq \mathbb{P}(\bigcup_{e,f} \mathcal{E}_{e,f}) \leq \sum_{e,f} \mathbb{P}(\mathcal{E}_{e,f})$ by the union bound.
  (i) If $|e \cap f| \neq 1$ then $\mathcal{E}_{e,f} = \emptyset$.
  (iii) if $|e \cap f| = 1$ then $\mathbb{P}[\mathcal{E}_{e,f}] = \frac{(k-1)! \cdot (k-1)!}{(2k-1)!}$. This can be seen by conditioning on the position of $e \cup f$ within the ordering and on the ordering of the vertices in $[n] \setminus (e \cup f)$. The number of orderings with any such condition is $(2k-1)!$. Then for $\mathcal{E}_{e,f}$ to happen we must first place the $k-1$ vertices in $e \setminus f$ in any order, then the vertex in $e \cap f$, and then the $k-1$ vertices of $f \setminus e$ in some order. This accounts for $(k-1)! \cdot (k-1)!$ possibilities.
- Estimation: Using the Stirling approximation: $\mathbb{P}(\mathcal{E}_{e,f}) \approx \sqrt{\frac{4\pi}{k}} 2^{-2k}$

  By the union bound, $\mathbb{P}(\text{algorithm fails}) \leq \sum_{e,f} \mathbb{P}(\mathcal{E}_{e,f}) \leq (1+o(1))m^2 \sqrt{\frac{4\pi}{k}} 2^{-2k} < 1$ provided

  $$m < \left( \frac{1}{\sqrt[4]{4\pi}} - o(1) \right) k^{\frac{1}{4}} 2^k.$$
- Conclusion: when $m = ck^{1/4}2^k$ the algorithm succeeds with positive probability, which implies that **there exists** a proper 2-coloring of $\mathcal{H}$.

$\square$

Now we will present an even better bound, which is a modification of Pluhár's proof by two Polish students(!) and reproves the bound of Radhakrishnan and Srinivasan (the best known lower bound).

**Theorem 0.2** (Cherkashin-Kozik, 2014). $m_B(k) \geq c\sqrt{\frac{k}{\ln k}} 2^k$.

*Proof (idea).* • Idea: randomised greedy algorithm as before. Twist: new "randomness".

- Random ordering:
  - For each vertex $v$, sample a real(!) number $x_v$ between 0 and 1, independently.
  - Order vertices in increasing order of $x_v$: $v$ before $w \iff x_v < x_w$.
  - In this order, run the greedy algorithm as described above.
- Comparison: This is exaclty the same probability space as in the Pluhár proof: the order of the vertices is chosen uniformly from all ordering. The continuous setting just makes the analysis more convenient. (not strictly necessary).

- Analysis: The algorithm is the same, so we must have some new trick to the analysis. In the Pluhár proof we had a bad event for every pair of edges that intersect in one vertex in case this vertex comes in the middle among the vertices of the two edges. We had no idea how many such pairs are in the hypergraph, so in our proof we assumed that all of them are like that. Now we reduce this by considering separately those edges whose last vertex comes unnaturally too far to the left or its first vertex unnaturally too far to the right. This is an unlikely event, as all vertices are "expected" to come around $\frac{1}{2}$. Coming too far to the left/right is an unlikely event for an edge, so we take care of these edges by introducing two new "bad events" for each edges. For the rest of the pairs of the edges however the common vertex *must* come closely around $\frac{1}{2}$, also unlikely.
- In Pluhár's proof: the pairs involving an unnaturally placed edge got counted multiple times.
- This proof: here they get counted once, yielding a better bound.
- Details: **HW**

$\square$

- Worst case scenario: Single vertex intersection
  - From these proofs it should be clear that edges intersecting in one vertex cause trouble.
  - In the HW it was proven that if there is no such pair of edges, then the graph is always 2-colourable.
  - having those edges makes $m_B(k)$ go from infinity to $O(k^2 2^k)$.
- Slightly paradoxical, but Cherkashin and Kozik also showed that the smallest non-2-colourable $k$-graph $\mathcal{H}$ with $|e \cap f| \leq 1$ for every two distinct edges $e, f \in \mathcal{H}$ must have *at least* $\Omega(\frac{4^k}{\ln^2(k)})$ edges.