Studia Scientiarum Mathematicarum Hungarica 30 (1995), 47–57

DISCREPANCY OF TREES

P. ERDŐS, Z. FÜREDI, M. LOEBL and V. T. SÓS

Abstract

We consider the question how large monochromatic part of a tree is present in any coloring of edges of a complete graph by two colors. It is proved that there exists a constant $c>0$ such that for any given tree T_n on n vertices with maximum degree Δ the following holds. An arbitrary coloring of the edges of K_n with 2 colors contains a copy of T_n such that at least $(n-1)/2+c(n-1-\Delta)$ edges of T_n get the same color.

1. Introduction, results

Discrepancy theory has originated from number theory. In the last two decades this subject has developed into an elaborate theory related also to geometry, probability theory, ergodic theory, computer science, combinatorics. The combinatorial setting of these problems proved to be a succesful approach. See the book of Beck and Chen [2], the chapter from the Handbook of Combinatorics [3], or [8].

One of the basic problems in combinatorial discrepancy theory is the following: Let $S=\{x_1,x_2,\dots,x_t\}$ be a finite set and $\mathcal{H}=\{A_1,\dots,A_m\}$ be a family of subsets of S. The goal is

(∗) to find a partition $S_1\cup S_2=S$, $S_1\cap S_2=\emptyset$ which splits each of the set in the family $\mathcal{H}$ as equally as possible.

A partition of S can be given by a function $\varphi : S\to\{1,2\}$. The discrepancy of $\mathcal{H}$ is defined by

$$\mathcal{D}(\mathcal{H}):=\min_{\varphi}\max_{A\in\mathcal{H}}\left||\varphi^{-1}(1)\cap A|-\frac{|A|}{2}\right|.$$

This measures, in supremum norm, how well the set S can be partitioned in the sense of (∗).

For a given $(S,\mathcal{H})$ we want to determine or estimate $\mathcal{D}(\mathcal{H})$. A large number of classical theorems in number theory, in geometry, in combinatorics can be formulated in this language. Here we consider the special case when

---

1991 *Mathematics Subject Classification.* Primary 05C05, 05D10, 05C55.

Key words and phrases. Discrepancy, Ramsey theory, extremal graphs, edge coloring, trees.

0081–6906/95/\$ 4.00

the underlying set S is the edge set, $E(K_n)$, of a complete graph and the family $\mathcal{H}$ is given by isomorphic copies of a given graph.

Let L be an arbitrary fixed graph. Our goal is to two-color the edges of K_n so that in each subgraph L^* isomorphic to L the edge-set $E(L^*)$ is two-colored as equally as possible. Let the two-coloring be given by $\varphi : E(K_n) \to \to \{1,2\}$. The *discrepancy* of L is defined by

$$\mathcal{D}_n(L,\varphi) := \max_{\substack{L^* \subseteqq K_n \\ L^* \sim L}} \left| |\varphi^{-1}(1) \cap E(L^*)| - \frac{|E(L^*)|}{2} \right|,$$

$$\mathcal{D}_n(L) := \min_{\varphi} \mathcal{D}_n(L,\varphi).$$

While Ramsey theory asks how large n should be so that any two-coloring of edges of K_n contains a monochromatic copy of a given graph L, discrepancy measures how large part of a graph L is present in any two-coloring. The case $L = K_t$ was investigated by Erdős and Spencer [5].

In this paper we consider the case when L is a tree T_n on n vertices. Put $\mathcal{D}(T_n) = \mathcal{D}_n(T_n)$.

Let S_n and P_n denote the star and the path on n vertices, respectively. It is obvious, that

$$\mathcal{D}(S_n) = \begin{cases} 0 & \text{for } n = 4k+1, \\ 1/2 & \text{for } n = 2k, \\ 1 & \text{for } n = 4k+3. \end{cases} \tag{1}$$

It is also easy to see that

$$\mathcal{D}(P_n) = \frac{1}{6}n + O(1).$$

This follows from a theorem of Gerencsér and Gyárfás [6] stating

$$R(P_k) = \lfloor (3k+1)/2 \rfloor, \tag{2}$$

where $R(L)$ denotes the *Ramsey number* of the graph L.

In general, $R(L_1, L_2)$ denotes the minimum integer n such that the following holds: for each coloring of the edge-set of $E(K_n)$ with the colors $\{1,2\}$ one can find either a copy of L_1 of color 1 or a copy of L_2 consisting of edges of color 2; finally $R(L) := R(L,L)$.

Which are the basic relevant properties of T_n determining whether $\mathcal{D}(T_n)$ is small or large?

Let $\Delta(L)$ denote the maximal degree in L. A set $C \subseteqq V(L)$ is called a *vertex cover* if each edge $e \in E(L)$ has at least one endpoint in C. Let $\tau(L_n)$ denote the minimum size of a vertex cover.

Here we prove that the order of magnitude of $\mathcal{D}(T_n)$ depends on $\Delta(T_n)$ and $\tau(T_n)$.

THEOREM 1.1. *Suppose that* $\Delta(T_n) \geqq 0.8n$. *Then* $\mathcal{D}(T_n) \geqq (n-1-\Delta)/6$.

For even n considering a two-coloring of $E(K_n)$ such that every color induces an $n/2$-regular graph, one sees that $\mathcal{D}(T_n) \leqq n-1-\Delta$.

THEOREM 1.2. *Suppose* $n > m_0$ *and* $\Delta(T_n) < 0.8n$. *Then* $\mathcal{D}(T_n) > n10^{-3}$.

Here the value of m_0 comes from Corollary 2.8.

The next theorem describes a class of trees having discrepancies as large as possible, $n/2 - o(n)$ (if $\max(\Delta(T_n), \tau(T_n)) = o(n)$).

THEOREM 1.3. *If* $\Delta(T_n), \tau(T_n) \leqq k \leqq n/8$, *then* $\mathcal{D}(T_n) \geqq (n/2) - 4k$.

Color red a complete subgraph of size $n-(k/2)$ and blue the rest of the edges of K_n. Then the largest monochromatic part of a tree with $\tau(T) = k$ does not have more than $n-(k/2)$ edges. Hence $\mathcal{D}(T_n) \leqq n/2 - k/2$.

2. Conjectures, problems, lemmata

The proofs of the theorems above are closely related to extremal and Ramsey problems on trees. Here a new type of extremal problems arose, where the lower bound on the number of edges (in Turán type problems) is replaced by a lower bound on the number of vertices with high degrees.

CONJECTURE 2.1 ($n/2$–$n/2$–$n/2$ conjecture). *Let* G *be a graph with* n *vertices and let at least* $n/2$ *of them have degree at least* $n/2$. *Then* G *contains any tree on at most* $n/2$ *vertices.*

M. Ajtai, J. Komlós and E. Szemerédi [1] proved the following approximate version.

THEOREM 2.2 (Ajtai, Komlós and Szemerédi [1]). *For every* $\eta > 0$ *there is a threshold* $n_0 = n_0(\eta)$ *such that the following statement holds for all* $n \geqq n_0$: *if* G *is a graph on* n *vertices, and at least* $(1+\eta)\frac{n}{2}$ *vertices have degrees at least* $(1+\eta)\frac{n}{2}$, *then* G *contains, as subgraphs, all trees with at most* $\frac{n}{2}$ *edges.*

J. Komlós and V. T. Sós extended Conjecture 2.1 for trees of any size.

CONJECTURE 2.3. *If* G *is a graph on* n *vertices and more than* $n/2$ *vertices have degrees greater than or equal to* k, *then* G *contains, as subgraphs, all trees with* k *edges.*

J. Komlós announced proving an approximate version of Conjecture 2.3, too.

THEOREM 2.4 (Komlós [7]). *For every $\eta > 0$ there is a threshold n_0 $= n_0(\eta)$ such that the following statement holds for all $n \geqq n_0$: if G is a gra on n vertices and at least $(1+\eta)\frac{n}{2}$ vertices have degrees at least $(1+\eta)k$ th G contains all trees with at most k edges.*

A weaker form of Theorem 2.4 which we will need follows analogously the proof of Theorem 2.2 [1].

THEOREM 2.5. *For every $\eta > 0$ there is a threshold $n_0 = n_0(\eta)$ such th the following statement holds for all $\varepsilon \geqq 0$ and $n \geqq n_0$: if G is a graph on vertices and at least $(1+\eta)\frac{n}{2}$ vertices have degrees at least $(1-\varepsilon+\eta)\frac{n}{2}$ th G contains all trees with at most $(1-3\varepsilon)\frac{n}{2}$ edges.*

SKETCH OF PROOF OF THEOREM 2.5. The proof goes in the same w as the proof of Theorem 2.2 in [1], with only one change: a combinator Lemma 6 of [1] is replaced by a lemma below proving a weaker proper (from weaker assumptions) than the original Lemma 6.

LEMMA 2.6. *Let H be a graph on N vertices, and let U be the set vertices of degree greater than $(1-\eta)\frac{N}{2}$. If $|U| \geqq \frac{N}{2}+1$ then there are t vertices x, $y \in U$ and a (partial) matching M in H such that*

x and y are adjacent,

M covers at least $(1-3\eta)\frac{N}{2}-1$ neighbors of both x and y.

PROOF OF LEMMA 2.6. First observe that at least two vertices of are joined by an edge of H. We will use the Gallai–Edmonds decompositi (GED). Let A be the set of vertices of H omitted by at least one maximu matching of H, let B the set of vertices of $H-A$ which have neighbors A and let C be the set of remaining vertices of H. GED Theorem asse that the connectivity components of $H-A$ are hypomatchable (a graph is called hypomatchable if $G-v$ has a perfect matching for each vertex of G), the connectivity components of $H-C$ have a perfect matching a any maximum matching of H covers completely B from A.

If a component of $H-B$ has two adjacent vertices of U then Lemma 2 follows. Hence U forms an independent set in each component of $H-$ Let α denote the size of a maximum independent set. However, $\alpha(C) <$ for any hypomatchable C with more than one vertex and $\alpha(C) \leqq \frac{|C|}{2}$ for a C with a perfect matching. Since $|U| > \frac{|V(H)|}{2}$, there is a hypomatchab component C of $H-B$ consisting of exactly one vertex which moreov belongs to U. Hence $|B| \geqq (1-\eta)\frac{n}{2}$ and by GED Theorem H has a matchi which covers at least $n-\eta n$ vertices.

Hence the Lemma 2.6 is proved and Theorem 2.5 then follows analogous as Theorem 2.2 in [1].

Using Theorem 2.5, it is not difficult to prove a Ramsey type resu which will provide a basic tool in further considerations.

THEOREM 2.7. *For every $\varepsilon > 0$ there is a threshold $m_0 = m_0(\varepsilon)$ such that the following statement holds for all $n \geqq m_0$: if G is a graph on n vertices and T is a tree on at most $(1-\varepsilon)\frac{n}{2}$ vertices, then G or complement of G contains T.*

PROOF. Let $m_0 = \left(1 - \frac{2}{9}\varepsilon\right)^{-1} n_0\left(\frac{\varepsilon}{9}\right)$, where $n_0(\cdot)$ comes from Theorem 2.5. Let $\eta = \frac{\varepsilon}{9}$ and $\varepsilon' = \frac{\varepsilon}{3}$. If G satisfies the assumptions of Theorem 2.5 for ε' and η then Theorem 2.7 follows, otherwise complement of G has at least $(1-\eta)\frac{n}{2}$ vertices of degree at least $n - (1 - \varepsilon' + \eta)\frac{n}{2} = \frac{n}{2}(1 + \varepsilon' - \eta)$. Denote by S the set of these vertices. Let G' be a graph obtained from $\overline{G}$ by deleting $2\eta n$ vertices from $V(\overline{G}) - S$. $|V(G')| = n' = (1-2\eta)n$. Now, at least $|S| \geqq (1-\eta)\frac{n}{2} \geqq (1+\eta)\frac{n'}{2}$ vertices of G' have degree at least $(1 + \varepsilon' - 5\eta)\frac{n}{2} \geqq$ $\geqq (1 - \varepsilon' + \eta)\frac{n'}{2}$. Since $n' \geqq n_0(\eta)$ we may apply Theorem 2.5 to G' and get that G' and hence also $\overline{G}$ has all trees on $(1-\varepsilon)\frac{n}{2}$ vertices. □

We will use only the following weaker version.

COROLLARY 2.8. *For $n > m_0$ the following holds. Every tree on at most $(\frac{1}{2} - 10^{-3})n$ vertices is contained in either G_n or in $\overline{G}_n$.*

Theorem 2.7 states that $R(T_k) \leqq 2k + o(k)$ as $k \to \infty$. Here we formulate the

CONJECTURE 2.9. *Let T_a and T_b be trees on a and b vertices, respectively, and let G be a graph on $a+b-2$ vertices. Then either G contains T_a or $\overline{G}$ contains T_b. Especially, $R(T_k) \leqq 2k - 2$.*

We think that even more is true.

CONJECTURE 2.10. *There is a $c > 0$ such that $R(T_k) < (2-c)k + c\Delta$.*

We conclude this section by an easy observation.

LEMMA 2.11. *Let M_a be a star-forest on $a \geqq 2$ vertices and consider an arbitrary two-coloring of the complete graph, $E(K_n) = E(G_1) \cup E(G_2)$. If G_1 does not contain a monochromatic copy of M_a then there is a subset $A \subseteqq$ $\subseteqq V(K_n)$ such that every vertex in A has more than $n - a$ G_2-neighbors in A. Consequently, $R(M_a, T_b) \leqq a + b - 2$.*

PROOF. If M consists of only one star, then the statement is trivial with $A = V(K_n)$. Otherwise, one can use induction on the number of stars in M.

If the degree of each vertex of the subgraph of G_2 induced on A is at least $n - a + 1$, then G_2 has every tree on $n - a + 2$ vertices. □

3. The case of large maximum degree

In this section we prove Theorem 1.1. Consider an arbitrary two coloring, φ, of the edge-set of the complete graph using the colors $\{1, 2\}$. Let T be an n vertex tree with $\Delta(T) \geqq 0.8n$. Suppose, on the contrary, that $\mathcal{D}_\varphi(T) = : x < (n-1-\Delta)/6$. Then $x < (n-1)/30$. Let S_1 be a monochromatic star of K_n of maximum number of vertices. Denote its vertex set by A_1, let $A_2 := V(K_n) - A_1$, and $|A_1| - 1 = (n-1)/2 + m$. Here $m \leqq x$. We may suppose that the edges of S are colored by the color 1.

Let M be the forest having $(n-1-\Delta)$ edges obtained from T by deleting the edges adjacent to a vertex of maximum degree. M has a subforest consisting of vertex disjoint stars and containing at least half of its edges. Let M_1, M_2 be star-forests contained in M of sizes $|E(M_1)| = x - m + 1$ and $|E(M_2)| = 3x - m + 1$. As the vertex of maximum degree of T is adjacent to at least $0.6(n-1)$ vertices of degree 1, T contains a vertex-disjoint copy of a star T_i and the star-forest M_i such that their total number of edges is $(n-1)/2 + x + 1$. (This is, indeed, a special case of Lemma 4.1.)

There is no monochromatic copy of M_1 in A_2 in color 1, otherwise together with S_1 it would form a too large monochromatic part of a copy of T. Hence Lemma 2.11 implies that there exists an $A_2' \subsetneqq A_2$ such that every degree in color 2 in A_2' is at least $|A_2| - 2(x - m + 1) + 1$. As the maximum degree in color 2 is at most $|A_1| - 1$ we obtain that every vertex of A_2' is joined to at most $2x + 1$ vertices from A_1 using edges of color 2.

We also obtain that there is a star S_2 of at least $|A_2| - 2(x - m) + 1$ edges of color 2 contained in A_2. Thus, repeating the previous argument, A_1 does not contain a copy of M_2 of color 2. Hence Lemma 2.11 implies that there exists an $A_1' \subsetneqq A_1$ such that every vertex in A_1' has degree in color 1 at least $|A_1| - 2(3x - m + 1) + 1$. We obtain that every vertex of A_1' is joined to at most $(6x - 2m)$ vertices of A_2' using edges of color 1.

Altogether, considering the complete bipartite graph with parts A_1' and A_2' we get that

$$2x + 1 + (6x - 2m) \geqq \min(|A_1'|, |A_2'|) \geqq (n-1)/2 + 1 + m - (6x - 2m + 1) + 1 .$$

This yields $x \geqq (n-1)/28$, a contradiction. □

4. How to cut a tree

In this section we collect some technical lemmata about tree decompositions we are going to use in the next section for the proof of our main result, Theorem 1.2. As we are providing an asymptotic only, for simplicity, from now on in this and the next sections, we suppose that n is even.

LEMMA 4.1. *Let T be a tree on n vertices and let $\Delta(T) < 0.8n$. Then there is a subtree T' on $n/2$ vertices and a subgraph M of T such that the following properties hold.*

(1) *M is star-forest of at least $(n-1)/16$ edges;*
(2) *M is vertex-disjoint to T'.*

PROOF. If there is a cut edge, e, of T such that the deletion of e results two trees on $n/2$-$n/2$ vertices, then we are done. Otherwise, T has a (unique) vertex, v with the following property: considering the edges $vv_1, vv_2, \dots, vv_t$ adjacent to v and the subtrees $T_1, \dots, T_t$, obtained after deleting all of these edges ($v \notin T_i$, $v_i \in T_i$), $s_i = |V(T_i)|$, we get that $s_1 \leqq s_2 \leqq \dots \leqq s_t < n/2$, (and, of course, $\sum s_i = n-1$). We have that $t \geqq 3$. Let j be defined by

$$1 + s_1 + \dots + s_{j-1} < n/2 \leqq 1 + s_1 + \dots + s_j.$$

Here $j < t$ (because $s_j < n/2$). Then T' can be any subtree of $v + T_1 + \dots + T_j$ on $n/2$ vertices. Define M' as the forest $T_{j+1} + \dots + T_t$. We claim that M' has at least $(n-1)/8$ edges. Indeed, if $s_j = 1$, then T' is a star and M' has at least $n - 1 - \Delta$ edges. Otherwise, for $s_j \geqq 2$ we have that

$$|E(M')| \geqq \sum_{i>j} (s_i - 1) \geqq \sum_{i>j} (s_j/2).$$

Here $\sum_{j>i} s_i > (n-1)/4$, because otherwise $s_j > (n-1)/4$ follows, and this again implies $(n-1)/4 < s_j \leqq s_{j+1}$. Finally, every forest contains a star-forest consisting of at least half of its edges, so there is an $M \subseteqq M'$ of size at least $(n-1)/16$. □

Considering the decomposition, $v + T_1 + \dots + T_{j-1}$, $v + T_j$, $v + T_{j+1} + \dots + T_t$ in the proof of Lemma 4.1 we obtain the following statement.

LEMMA 4.2. *The edge set of an arbitrary tree T can be partitioned into at most 3 trees each of sizes at most $|V(T)|/2$.*

Let W be the set of all neighbors of leaves of T. For each $w \in W$ choose a neighboring vertex of degree 1, we get the set W', $|W| = |W'|$. Applying Lemma 4.2 for the tree $T - W'$ one gets the following

COROLLARY 4.3. *$T - W'$ has a subtree T' on $\frac{n}{2}$ vertices, which contains at least $\frac{1}{3}|W|$ vertices of W.*

Let P be the set of pendant edges. Deleting $\deg(x) - 2$ edges from each vertex x of degree at least 3 one gets a subforest which is a path-forest, i.e., we obtain the following

LEMMA 4.4. *T_n has a subforest T' of at least $n - |P|$ edges consisting of vertex-disjoint paths, edges and isolated vertices.*

5. Proof of Theorem 1.2

Let $n > m_0$, and let T be a tree on n vertices satisfying $\Delta(T) < 0.8n$. Let φ be a two-coloring of the edges of K_n, and suppose, on the contrary, that $\mathcal{D}_\varphi(T) < n/258 =: l$.

CLAIM 5.1. *The vertices of* K_n *may be partitioned by* $V(K_n)=A_1\dot\cup A_2$, $|A_1|=|A_2|=\frac{n}{2}$ *so that both graphs*

$$G_i=\{e\subset A_i : e \text{ has color } i\},\quad i=1,2$$

contain all trees on $\frac{n}{2}-8l$ *vertices. Moreover, there are sets* $B_i\subseteqq A_i$ *such that the minimum degree of the restriction of* G_i *to* B_i *is at least* $n-8l$.

PROOF. Let T_1' be a subtree on $\frac{n}{2}-l$ vertices provided by Lemma 4.1. Let T_2' be a subtree of T_1' of $n/2-3l$ edges. Also let M_1 and M_2 be star-forest contained in T vertex-disjoint to T_1' and T_2', respectively, of sizes $|E(M_1)|= =2l$, $|E(M_2)|=4l$.

We use Corollary 2.8 to find a monochromatic copy of T_1', say color 1. Let A_1' be formed by the vertices of this copy of T_1'. There is no copy of M_1 of color 1 vertex disjoint to A_1', otherwise we obtain $\mathcal{D}_\varphi(T)\geqq l$. By Lemma 2.11 we have that $V(K_n)-A_1'$ contains a copy of T_2' of color 2. Then define the sets A_1, A_2 such that $V(T_i')\subseteqq A_i$, $|A_i|=n/2$, $A_1\cup A_2=V(K_n)$. The set A_i does not contain a copy of M_{3-i} of color $3-i$. Hence Lemma 2.11 yields that A_i contains a set B_i satisfying the requirements and B_i contains all trees of color i of sizes at most $n/2-8l$. □

To finish the proof of Theorem 1.2 we distinguish three cases.

1. $|W|\geqq 54l$, where W is the set of all neighbors of leaves of T. By Corollary 4.3 there is a subtree T' on $\frac{n}{2}$ vertices and a matching M such that each edge of M intersects T' in one vertex and $|M|\geqq 18l$. In each A_i, $i=1,2$ take a copy of T_i' with at least $\frac{n}{2}-8l$ edges of color i. Between $(M\cap T_i')\cap A_i$, $i=1,2$ there must be a monochromatic matching of at least $9l$ edges. This together with the corresponding copy of T' has at least $\frac{n}{2}+l$ edges of the same color. This finishes Case 1.

2. $|P|<n/4-(3/2)l$, where P is the set of pendant edges. By Lemma 4.4, T contains a path-forest, T', of at least $n-|P|$ edges. We apply (2) that K_n contains a monochromatic path, H, of at least $(2/3)(n-1)$ edges. We can cover at least 2/3 of the edges of T' by H and conclude that T has a monochromatic part of at least $n/2+l$ edges. This finishes Case 2.

3. If neither Case 1 nor Case 2 take place then let T' be a subtree of $(n/2)-|W|-8l$ edges on $(n/2)-8l$ vertices obtained from T by deleting edges in the following 3 steps. Let P' be a set of $(n/4)-(3/2)l$ pendant edges, delete these from T. Second, delete $|W|-1$ edges such that the rest of the tree consists of $|W|$ components each component having exactly one vertex from W. Finally, trim leaves off these components to get the desired size such that we never cut off a vertex of W.

Without loss of generality we may assume that A_1 has a set N of $\frac{n}{4}$ vertices such that each of them is incident with at least $\frac{n}{4}$ edges of color 1 going to A_2. Fix a copy of T' in B_1 such that the vertices of W all come from $B_1\cap N$. The edges of P' can be added to T' from the color 1 edges between A_1 and A_2. We found a subgraph of T with at least $\frac{3n}{4} - 63.5l$ edges of color 1. This finishes Case 3, thus Theorem 1.2 is proved.

6. Proof of Theorem 1.3

Let T be a tree on n vertices and consider an arbitrary two-coloring of the edges of K_n using colors red and blue. We claim that K_n contains a subforest of T of at least $n-4k$ edges of the same color, consisting of vertex-disjoint stars.

Let T^* be a maximum star-forest of T. T^* has at least $n-\tau$ edges. Let T' be a maximum monochromatic subgraph of T^*. If T' has at least $n-4k$ edges we are done. In the rest of the proof we assume that T' has less than $n-3k$ vertices.

Let us assume that the color of T' is red. Let x be a vertex of T' of degree 1. There are less than k red edges going from x to vertices out of T' in T, otherwise T' may be improved by replacing the edge incident with x by the red star of k edges rooted in x, whose leaves do not belong to T'. This new system of red stars contains a subgraph of T^* which is bigger than T'. Similar argument shows that red stars of $K_n - V(T')$ have at most $(k-1)$ edges. Let $W = V(K_n) - V(T')$ and let $W = W_1 \cup W_2$ be a partition of W such that $|W_1| = 3k$. We will construct a big blue subgraph of T^*. Its stars will be rooted in W_1 and leaves will be in $V(K_n) - W_1$. Let T'' denote the current part of this blue subgraph which we have already constructed. We enlarge T'' as follows. If there are at least $2k$ vertices of $M = W_2 \cup \{j;\ j$ is a vertex of T' of degree $1\}$ uncovered by T'' then observe that at least one vertex of $W_1 - T''$ is incident by blue edges with at least k vertices of M. Thus we enlarge T'' by adding this star to it. If less than $2k$ vertices of M are uncovered by T'' then we stop. In the end T'' has at least $n-3k$ vertices out of W_1, hence it has at least $n-4k$ edges. Hence Theorem 1.3 is proved.

7. Further problems and generalizations

Above the special case was considered when $E(K_n)$ was two-colored, and we investigated how large monochromatic portion of a given tree T_n must be contained in it. Here we give a list of some possible generalizations.

(1) Instead of K_n we can consider other sequence of underlying graphs, e.g., the complete bipartite $K_{n,n}$, t-partite graphs $K_{n,n,\dots,n}$;

(2) Instead of copies of a T_n some other family of graphs, even with different sizes can be investigated;

(3) Two coloring can be replaced by r-coloring;

(4) Instead of the measuring the disrepancy in supremum norm it is interesting to consider the average, e.g., the l_2 norm;

(5) Instead of considering the maximum distance from the evenly colored subgraphs (when the goal was to approach a $(1/2, 1/2)$ coloring) to consider for a given $\alpha \in (0,1)$ the discrepancy from an $(\alpha, 1-\alpha)$ coloring. Some applications lead these kind of questions, eventually α depends on n, $\alpha = \alpha(n)$;

Finally we mention two further problems.

1. A general method in discrepancy theory is to obtain an estimation from the discrepancy of the random coloring. One of the first problems is to decide when the random coloring yields the optimal or nearly optimal solutions. It is easy to see that when $|E(L)| = \omega(n) n \log n$ with $\omega(n) \to \infty$, for $n \to \infty$, then already the random coloring φ gives $\mathcal{D}_n(L, \varphi) = o(|E(L)|)$.

2. In our case (the case of spanning trees) the bounds on the discrepancy are in terms of the maximum degree, Δ, and the covering number, τ. It would be interesting to see what other graph parameters or structural properties of the sample graphs (and the underlying graphs) influence the discrepancy. For example, if the tree T_n has two vertices of degree $n/2$ (it is called a *broom*), then $\mathcal{D}(T_n) = n/4 + O(1)$. (This was also proved by Bondy [4].)

ACKNOWLEDGEMENT. The authors would like to thank János Komlós, Imre Ruzsa and Mikkel Thorup for helpful discussions. This research of the second author was partially supported by an NSF Grant. The research of the fourth author was partially supported by the Hungarian National Foundation for Scientific Research Grant No. 1909.

REFERENCES

[1] AJTAI, M., KOMLÓS, J. and SZEMERÉDI, E., On a conjecture of Loebl, *Proc. 7th International Conference on Graph Theory*, Kalamazoo, Michigan, 1993.

[2] BECK, J. and CHEN, W. L., *Irregularities of distribution*, Cambridge Tracts in Mathematics, 89, Cambridge Univ. Press, Cambridge–New York, 1987. *MR* **88m**: 11061

[3] BECK, J. and SÓS, V. T., Discrepancy theory, *Handbook of combinatorics*, Springer, Berlin–New York, 1994.

[4] BONDY, A., Personal communication.

[5] ERDŐS, P. and SPENCER, J., Imbalances in k-colorations, *Networks* **1** (1971/72), 379–385. *MR* **45** #8573

[6] GERENCSÉR, L. and GYÁRFÁS, A., On Ramsey-type problems, *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.* **10** (1967), 167–170. *MR* **39** #1351

[7] KOMLÓS, J., Personal communication, 1993.

[8] SÓS, V. T., Irregularities of partitions: Ramsey theory, uniform distribution, *Surveys in combinatorics* (Southampton, 1983), London Math. Soc. Lecture Note Ser., 82, Cambridge Univ. Press, Cambridge–New York, 1983, 201–246. *MR* **85h**:05012

(Received February 14, 1994)

P. Erdős and V. T. Sós
MTA MATEMATIKAI KUTATÓINTÉZETE
POSTAFIÓK 127
H–1364 BUDAPEST
HUNGARY

Z. Füredi
MTA MATEMATIKAI KUTATÓINTÉZETE
POSTAFIÓK 127
H–1364 BUDAPEST
HUNGARY

and

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ILLINOIS
URBANA, IL 61801–2917
U.S.A

M. Loebl
KATEDRA APLIKOVANÉ MATEMATIKY
MATEMATICKO–FYZIKÁLNI FAKULTA
UNIVERZITA KARLOVA
MALOSTRANSKÉ NÁM. 25
CZ–118 00 PRAHA 1
CZECH REPUBLIC