Electronic Colloquium on Computational Complexity, Report No. 21 (2005)

DISPROVING THE SINGLE LEVEL CONJECTURE

STASYS JUKNA[†][‡][§]

Abstract. We consider the minimal number of AND and OR gates in monotone circuits for quadratic boolean functions, i.e. disjunctions of length-2 monomials. The single level conjecture claims that monotone single level circuits, i.e. circuits which have only one level of AND gates, for quadratic functions are not much larger than arbitrary monotone circuits. In this paper we disprove the conjecture: there are quadratic functions in n variables whose monotone circuits have linear size whereas their monotone single level circuits require size $\Omega(n^{2-\varepsilon})$.

Key words. Monotone circuits, quadratic functions, clique covering number, Kneser graph, Sylvester graph.

AMS subject classifications. 05C35, 05C60, 68Q17, 68R10, 94C30

1. Preface.

Every graph $G = (V, E)$ defines a natural (boolean) *quadratic function*

$$f_G(X) = \bigvee_{uv \in E} x_u x_v.$$

We consider the complexity of computing such functions by monotone circuits over the standard monotone basis $\{\vee, \wedge, 0, 1\}$ of fanin-2 AND and OR gates. Single level circuits are circuits where every path from an input to the output gate contains at most one AND gate.

SINGLE LEVEL CONJECTURE: *For quadratic functions single level circuits are almost as powerful as unrestricted ones.*

This conjecture—first explicitly named the "single level conjecture" by Lenz and Wegener in [12]—was considered by several authors, [11, 3, 4, 13, 12, 2] among others. A strong support for this conjecture was given by Mirwald and Schnorr [13]: if we consider circuits over the basis $\{\oplus, \wedge, 0, 1\}$ computing quadratic forms $f_G^{\oplus}(X) = \bigoplus_{uv \in E} x_u x_v$ over GF(2) and if we count only AND gates, then *every* optimal (with respect to the number of AND gates) circuit is a single level circuit.

In this paper we show that in the basis $\{\vee, \wedge, 0, 1\}$ the single level conjecture is not even near to the truth: there are quadratic functions in n variables whose monotone circuits have linear size whereas their monotone single level circuits require size $\Omega(n^2/\log^3 n)$. Similar gaps are shown for the multiplicative complexity (when we count only AND gates) as well as for boolean formulas. Finally, we give an indication that the single level conjecture should also fail in the case of circuits with unbounded fanin gates.

[†] *Current address*: Universität Frankfurt, Institut für Informatik Robert-Mayer-Str. 11-15, D-60054 Frankfurt, Germany. Email: jukna@thi.informatik.uni-frankfurt.de

[‡] Institute of Mathematics and Informatics, Akademijos 4, LT-08663 Vilnius, Lithuania.

[§] Research supported by a DFG grant SCHN 503/2-2.

ISSN 1433-8092

2. Introduction and results. Given a graph $G = (V, E)$ we associate to each of its vertices v a boolean variable x_v, and let $X = \{x_v : v \in V\}$. A *non-edge* is a pair uv of non-adjacent vertices. If $G \subseteq U \times W$ is a bipartite graph with parts (or color classes) U and W, then a non-edge is a pair uv of non-adjacent vertices with u and v belonging to different color classes; hence, in bipartite case, pairs of vertices of the same color are neither edges nor non-edges. For convenience, we often look at boolean functions $f(X)$ as accepting/rejecting subsets of vertices $S \subseteq V$: the function accepts/rejects a subset of vertices if it accepts/rejects the incidence vector of this subset.

A boolean function (or a circuit) *represents* a graph if it accepts all edges and rejects all non-edges; on other inputs the function can take arbitrary values.

The reason why (even monotone!) circuit complexity of graph representation is interesting is the following. Every bipartite $n \times n$ graph $G \subseteq U \times W$ with $n = 2^m$ and $U = W = \{0,1\}^m$ gives us a boolean function f (the characteristic function of G) in $2m$ variables such that $f(uv) = 1$ iff $uv \in G$. If we have a non-monotone circuit for f then it is possible to replace its input literals by boolean sums (ORs of variables) so that the resulting monotone circuit represents G (see [16, 9]). Since, as shown in [16] (see Lemma 13 below) $4m = 4\log n$ boolean sums can be simultaneously computed by a monotone circuits of size $12n$, this implies that a lower bound $12n + n^{\varepsilon}$ on the size of *monotone* circuits representing G would yield a lower bound $n^{\varepsilon} = 2^{\varepsilon m}$ on the *non-monotone* circuit size of an explicit boolean function in $2m$ variables.

It is clear that the quadratic function $f_G(X) = \bigvee_{uv \in E} x_u x_v$ represents the graph $G = (V, E)$. Moreover, for some graphs G, f_G is the *only* monotone boolean function representing G (see Fact 2 below). It is therefore important to better understand the structure of circuits computing quadratic functions.

Mirwald and Schnorr [13] have investigated $C_{\mathrm{mult}}(f_G^{\oplus})$, the *multiplicative* complexity of (algebraic) quadratic forms, i.e. the minimal number of AND gates in a circuit over the basis $\{\oplus, \wedge, 0, 1\}$ for $f_G^{\oplus}(X) = \bigoplus_{uv \in E} x_u x_v$ and $C^1_{\mathrm{mult}}(f_G^{\oplus})$, the *multiplicative single level* complexity of $f_G^{\oplus}$. They proved that, with respect to the multiplicative complexity, the single level conjecture in this basis holds in the following strong sense.

Theorem 1 (Mirwald–Schnorr [13]). *$C^1_{\mathrm{mult}}(f_G^{\oplus}) = C_{\mathrm{mult}}(f_G^{\oplus})$ for every graph G. Moreover,* each *optimal* (*with respect to the number of AND gates*) *circuit for $f_G^{\oplus}$ is a single level circuit.*

Motivated by this result, Lenz and Wegener [12] considered the multiplicative complexity of quadratic forms f_G over the basis $\{\vee, \wedge, 0, 1\}$. Among many other results (including the algorithmic aspects of estimating the circuit complexity of quadratic functions) they observed that in this basis $C^1_{\mathrm{mult}}(f_G)$ is just the bipartite clique covering number $\mathrm{cc}(G)$ of G, i.e. the minimal number of bipartite complete subgraphs of G covering all edges of G. Hence, $C^1_{\mathrm{mult}}(f_G) \leq n$ for every n-vertex graph; in fact, $C^1_{\mathrm{mult}}(f_G) \leq n - \lfloor \log_2 n \rfloor - 1$ by a result of Tuza [19]. They also constructed a graph G such that $C_{\mathrm{mult}}(f_G) = 3$ but $C^1_{\mathrm{mult}}(f_G) = 4$, and asked how large can the gap

$$\mathrm{Gap}_{\mathrm{mult}}(G) = C^1_{\mathrm{mult}}(f_G)/C_{\mathrm{mult}}(f_G)$$

for n-vertex graphs be ([12], Problem 7). It turns out that this gap may be huge. We will show that the gap may be huge also in the case when we count all gates (not just AND gates) as well as in the case of formulas. We will also give an indication that the single level conjecture fails even if we allow *unbounded* fanin gates.

We will show the corresponding gaps for graphs G whose representation is not easier than the computation of f_G. If a function f represents a graph G, then it may (wrongly) accept some independent sets of G of size larger than two. The simplest way to exclude this possibility is to "kill off" all such independent sets by "saturating" the graph, i.e. by adding new edges. This is a well-known trick in the theory of boolean functions to obtain so-called *slice* functions (see, e.g. [23]).

By an *extension* of a bipartite graph $H \subseteq U \times W$ we mean any graph $G = (V, E)$ with $V = U \cup W$ and $E \cap (U \times W) = H$. Such an extension is *saturated* if the induced subgraphs of G on U as well as on W are complete graphs. A *complete star* in a graph with n vertices is a set of $n-1$ edges sharing one endpoint in common. If the graph is bipartite, then a complete star is a set of edges joining all vertices of one part with a fixed vertex of the other part. A graph is *star-free* if it contains no complete stars.

FACT 2. *Let G be a saturated extension of a bipartite star-free graph. Then f_G is the only monotone boolean function representing G.*

Proof. Let $G = (V, E)$ be a saturated extension of a bipartite graph $H \subseteq U \times W$. Suppose that H has no complete stars, and let f be a monotone boolean function representing G. Take an arbitrary subset $S \subseteq V$ of vertices. If $f_G(S) = 1$ then S contains both endpoints of some edge $uv \in E$. This edge must be accepted by f and, since f is monotone, $f(S) = 1$. If $f_G(S) = 0$ then S is an independent set of G. But the only independent sets in G are single vertices and non-edges of H. Hence, $f(S) = 0$ because f represents H and H contains no complete stars. □

A *perfect matching* of size n (or an *n to n matching*) is a bipartite $n \times n$ graph consisting of n vertex disjoint edges.

THEOREM 3. *If G is the saturated extension of a perfect matching of size n then $C_{\mathrm{mult}}(f_G) = O(\log n)$ but $C^1_{\mathrm{mult}}(f_G) \geq n/2$. Hence,* $\mathrm{Gap}_{\mathrm{mult}}(G) = \Omega(n/\log n)$.

This lower bound is not very far from the maximal possible because $C^1_{\mathrm{mult}}(G) \leq n$ for every graph G on n vertices. A better upper bound $\mathrm{Gap}_{\mathrm{mult}}(G) = O(n/\log\log n)$ was recently proved by Amano and Maruoka in [2].

The theorem itself is implicitly contained in [9] where it is shown that an n to n matching can be represented by a monotone CNF with $O(\log n)$ clauses. Here we state it explicitly. Theorem 3 was recently rediscovered in [2].

Recall that all these results concern the *multiplicative* complexity where we count only AND gates. The status of the single level conjecture in the case of *combinational* complexity—where we count *both* AND and OR gates—remained unclear. In this case very little was known on what the gap

$$\mathrm{Gap}(G) = C^1(f_G)/C(f_G)$$

really is (cf. Problem 6 in [12]). Krichevski [11] has proved that $\mathrm{Gap}(G) = 1$ for the complete graph $G = K_n$, even if negation is allowed as an operation. In the case of formulas, a graph with gap 8/7 was given by Bublitz [4]. Amano and Maruoka [2] have recently shown the gap of 29/28 for circuits computing *sets* of quadratic functions. However, even the existence of a (single) graph G with $\mathrm{Gap}(G) > 1$ was not known.

The following theorem disproves the single level conjecture in a strong sense.

THEOREM 4. *There exist n-vertex graphs G such that $C(f_G) = O(n)$ but $C^1(f_G) = \Omega(n^2/\log^3 n)$. Hence,* $\mathrm{Gap}(G) = \Omega(n/\log^3 n)$.

The graphs used in Theorem 4 are saturated extensions of Sylvester graphs, i.e. of bipartite graphs whose vertices are vectors in $\mathrm{GF}(2)^r$, and where two vertices are adjacent iff their scalar product over GF(2) is 1.

Next we consider the single level conjecture for formulas. Recall that a *formula* is a circuit where all gates have fanout 1; its *length* is the number of input gates. Let $L(f)$ and $L^1(f)$ denote the minimum length of a monotone (resp., in a monotone single level) formula computing f. Let

$$\mathrm{Gap}_{\mathrm{form}}(G) = L^1(f_G)/L(f_G).$$

As we already mentioned above, a graph G with $\mathrm{Gap}_{\mathrm{form}}(G) = 8/7$ was given by Bublitz [4]. However, it was open whether $\mathrm{Gap}_{\mathrm{form}}(G) = O(1)$ for all graphs G. That this is true for a large class of graphs follows from

THEOREM 5 ([9]). *If a graph $G = (V, E)$ has no triangles and no 4-cycles, then $L(f_G) \geq |E|/2$.*

Since, by the definition of f_G, $L^1(f_G) \leq 2|E|$ for very graph G, this implies that $\mathrm{Gap}_{\mathrm{form}}(G) \leq 4$ for a large class of graphs. Still, it turns out that also in the case of formulas, the single level conjecture is not even near to the truth.

A bipartite *Kneser* $n \times n$ graph is a bipartite graph $K \subseteq U \times W$ where U and W consist of all $n = 2^r$ subsets u of $\{1, \ldots, r\}$, and $uv \in K$ iff $u \cap v = \emptyset$.

THEOREM 6. *If G is the saturated extension of a bipartite Kneser $n \times n$ graph, then $L(f_G) = O(n \log n)$ but $L^1(f_G) \geq n^{1+c}$ for a constant $c > 0$. Hence,* $\mathrm{Gap}_{\mathrm{form}}(G) = n^{\Omega(1)}$.

So far we considered circuits with AND and OR gates of fanin 2. But what happens if we allow gates of *arbitrary* fanin—does the single level conjecture holds for such circuits? In this case the single level circuits are precisely the Σ_3 circuits. Recall that these circuits consist of unbounded fanin AND and OR gates which are organized in three levels: the bottom (next to the inputs) level consists of OR gates, the middle level consists of AND gates, and the top level consists of a single OR gate.

For a graph G, let $C_*(f_G)$ (resp., $L_*(f_G)$) be the minimum size of a monotone unbounded fanin circuit (resp., formula) computing f_G. Let also $C^1_*(f_G)$ and $L^1_*(f_G)$ denote the corresponding measures in a class of monotone Σ_3 circuits (i.e. the single level versions of these measures).

What are the gaps $\mathrm{Gap}^*(G) = C^1_*(f_G)/C_*(f_G)$ and $\mathrm{Gap}^*_{\mathrm{form}}(G) = L^1_*(f_G)/L_*(f_G)$?

The question is interesting because the presence of unbounded fanin gates may *exponentially* increase the power of single level circuits: there are n-vertex graphs G such that $C^1(f_G) = \Omega(n)$ but $C^1_*(f_G) = O(\log n)$ (we will show this in § 7).

The following result gives an (indirect) indication that the single level conjecture should fail also for unbounded fanin circuits or formulas.

THEOREM 7. *If the single level conjecture holds for unbounded fanin circuits or formulas then* $C_*(f_G) = \Omega(\sqrt{M}/d)$ *or* $L_*(f_G) = \Omega\left(M/d^2\right)$ *for every star-free graph with* M *edges and maximum degree* d.

The rest of the paper is organized as follows. In the next section we prove several auxiliary lemmas, relating the circuit complexity of quadratic functions f_G to some combinatorial characteristics of their graphs G. We then use these lemmas to prove Theorems 3, 4 and 6 in §§ 4-6. In § 7 we consider the single level conjecture in the case of circuits with *unbounded* fanin AND and OR gates and give an indication that the conjecture should fail also in this circuit model. We conclude with several problems.

3. Combinatorics of single level complexity. In this section we prove several auxiliary lemmas allowing us to get small upper bounds for circuit complexity of quadratic functions and large lower bounds on their single level complexity.

3.1. Upper bounds. A monotone CNF (*conjunctive normal form*) of *length* r is an AND of r *clauses*, each being an OR of variables; the *length* of a CNF is the number of clauses. CNFs of length at least three are simplest circuits violating the single level restriction. Thus, graphs represented by short CNFs may be good candidates to refute the single level conjecture. We will use such graphs in the proof of Theorems 3 and 6.

Let $\mathrm{cnf}(G)$ be the minimum length of a monotone CNF representing the graph G. As observed in [9], this number can be combinatorially described in terms of set-intersections. Say that a graph $G = (V, E)$ admits an *intersection representation of size* r if it is possible to associate with every vertex $u \in V$ a subset A_u of $\{1, \ldots, r\}$ so that $A_u \cap A_v = \emptyset$ if uv is an edge, and $A_u \cap A_v \neq \emptyset$ if uv is an non-edge of G. Let $\mathrm{int}(G)$ denote the smallest r for which G admits such a representation.

Erdős, Goodman and Pósa [7] observed that $\mathrm{int}(G)$ coincides with the clique covering number of $\overline{G}$, i.e. the minimum number of independent sets of G covering all non-edges of G. It turns out that this number also captures the length of CNFs representing G.

FACT 8 ([9]). *For every graph G,* $\mathrm{cnf}(G) = \mathrm{int}(G)$*.*

Proof. If a graph G can be represented by a CNF $\bigwedge_{i=1}^{r} \bigvee_{v\in S_i} x_v$, then the sets $A_u = \{i : u \notin S_i\}$ give the desired intersection representation of G. Conversely, having an intersection representation $\{A_u : u \in V\}$ of G, the CNF given above with $S_i = \{u : i \notin A_u\}$ represents the graph G. ∎

The following lemma shows that negation is (almost) powerless in the context of graph representation.

LEMMA 9. *Let H be a bipartite $n \times n$ graph. If H can be represented by a circuit of size L over the basis $\{\vee, \wedge, \neg\}$, then H can be represented by a monotone circuit of size at most $2L + O(n)$.*

Proof. The proof is a reminiscent of the proof that negation is powerless for slice functions (see, e.g. [23], p. 196).

Let F be a circuit of size L over the basis $\{\vee, \wedge, \neg\}$ representing a bipartite graph $H \subseteq U \times W$. Using DeMorgan rules we can transform this circuit to a circuit F' of size at most $2L$ such that negation is used only on inputs. We then replace each negated input $\overline{x}_u$ with $u \in U$ by a boolean sum $g_u = \bigvee_{v\in U\setminus\{u\}} x_v$, and replace each negated input $\overline{x}_w$ with $w \in W$ by a boolean sum $h_w = \bigvee_{v\in W\setminus\{w\}} x_v$. Since all these boolean sums can be simultaneously computed by a trivial circuit consisting of $O(n)$ OR gates (see, e.g. [23], p. 198 for a more general result), the size of the new circuit F_+ does not exceed $2L + O(n)$. Since the only difference of F_+ from the original circuit F is that negated inputs are replaced by boolean sums, it remains to show that on arcs $ab \in U \times W$ these sums take the same values as the corresponding inputs.

Take an arbitrary set $S = \{a, b\}$ with $a \in U$ and $b \in W$. The incidence vector of this set has precisely two 1's in positions a and b. Hence, $g_u(S) = 1$ iff $a \neq u$ iff $x_u(S) = 0$ iff $\overline{x}_u(S) = 1$. Similarly, $h_w(S) = 1$ iff $b \neq w$ iff $x_w(S) = 0$ iff $\overline{x}_w(S) = 1$. Hence, on edges and non-edges of H the functions g_u and h_w take the same values as the negated variables $\overline{x}_u$ and $\overline{x}_w$, implying that F_+ represents H. ∎

If a bipartite graph can be *represented* by a small circuit then the quadratic function of its saturated extension cannot require large circuits: it is enough to additionally compute two threshold-2 functions. We make this simple observation explicit in the following lemma.

LEMMA 10. *Let $H \subseteq U \times W$ be a bipartite $n \times n$ graph, G the saturated extension of H, and f a boolean function representing H. Then $f_G = (f \wedge g) \vee h$ where g is an AND of two monotone clauses and h is an OR of $O(\log n)$ monotone CNFs of length 2. Hence, $C_{\mathrm{mult}}(f_G) \leq C_{\mathrm{mult}}(f) + O(\log n)$, $C(f_G) \leq C(f) + O(n)$ and $L(f_G) \leq L(f) + O(n \log n)$. Moreover, if H is star-free then $f_G = f \vee h$.*

Proof. Let $g = \left(\bigvee_{u\in U} x_u\right) \wedge \left(\bigvee_{w\in W} x_w\right)$ and $h = T_2^U \vee T_2^W$ where $T_2^U(S) = 1$ iff $|S \cap U| \geq 2$, i.e. T_2^U is a threshold-2 function on n variables $\{x_u : u \in U\}$. Since the complete graph K_n can be covered by $m \leq \lceil \log n \rceil$ bipartite cliques, each of the functions T_2^U and T_2^W have the form $\bigvee_{i=1}^{m} \left(\bigvee_{u\in A_i} x_u\right) \wedge \left(\bigvee_{v\in B_i} x_v\right)$. Hence, h can be computed by an OR of $O(\log n)$ monotone CNFs of length 2. This immediately implies that $C_{\mathrm{mult}}(f_G) \le C_{\mathrm{mult}}(f) + O(\log n)$ and $L(f_G) \le L(f) + O(n \log n)$. That h can be computed by a monotone circuit with $O(n)$ gates follows from Lemma 13 stated below; this also follows from the fact that $C(T_k^n) \le kn + o(n)$ for any constant k ([8], see also [23], p. 152). It remains to show that $(f \wedge g) \vee h$ coincides with f_G.

If $f_G(S) = 1$ then S contains both endpoints of some edge $uv \in E$ of G. This edge must be accepted either by $f \wedge g$ (if $uv \in H$) or by h (if both u and v are in the same color class). Since both $f \wedge g$ and h are monotone, the function $(f \wedge g) \vee h$ accepts S.

If $f_G(S) = 0$ then S is an independent set of G, that is, S is either a single vertex or a non-edge of H. In both cases $h(S) = 0$ because none of the color classes can contain more than one vertex from S. Moreover, $g(S) = 0$ if S is a single vertex, and $f(S) = 0$ if S is a non-edge of H.

If H is star-free then the function f alone rejects all single vertices and non-edges of H, implying that in this case $f_G = f \vee h$. □

3.2. Lower bounds. A *bipartite clique covering* of G is a family of complete bipartite subgraphs $A_1 \times B_1, \ldots, A_t \times B_t$ of G such that every edge of G is an edge of at least one member of the family. The number t of subgraphs in such a covering is the *size* and the total number $\sum_{i=1}^{t}(|A_i| + |B_i|)$ of vertices is the *weight* of the covering.

Let $\mathrm{cc}(G)$ denote the minimum size and $\mathrm{cc_w}(G)$ the minimum weight of a bipartite clique covering of G. These measures were first studied by Erdős, Goodman and Pósa in [7], and now are the subject of extensive literature. In particular, it is known that the maximum of $\mathrm{cc}(G)$ over all n-vertex graphs is $n - \Theta(\log n)$ [19, 18], and that the maximum of $\mathrm{cc_w}(G)$ is $\Theta(n^2/\log n)$ [6, 4].

Let $w(G)$ be the minimum of $(a+b)/ab$ over all pairs $a, b \ge 1$ such that G contains a copy of $K_{a,b}$.

LEMMA 11. *For every graph G we have $C^1_{\mathrm{mult}}(f_G) = \mathrm{cc}(G)$ and*

$$L^1(f_G) = \mathrm{cc_w}(G) \ge w(G) \cdot |E|.$$

Moreover, if G is an extension of a bipartite graph H, then $\mathrm{cc}(G) \ge \mathrm{cc}(H)/2$ and $\mathrm{cc_w}(G) \ge \mathrm{cc_w}(H)$.

Proof. The equalities $C^1_{\mathrm{mult}}(f_G) = \mathrm{cc}(G)$ and $L^1(f_G) = \mathrm{cc_w}(G)$ follow immediately from a simple observation (made also in [3, 12]) that every single level circuit for f_G is of the form $\bigvee_{i=1}^{t} \left(\bigvee_{u \in A_i} x_u\right) \wedge \left(\bigvee_{v \in B_i} x_v\right)$ with $A_i \cap B_i = \emptyset$ for all $i = 1, \ldots, t$.

To prove that $\mathrm{cc_w}(G) \ge w(G) \cdot |E|$, let $E = A_1 \times B_1 \cup \cdots \cup A_t \times B_t$ be a bipartite clique covering of $G = (V, E)$ of minimal weight $\sum_{i=1}^{t}(|A_i| + |B_i|)$. Select subsets $E_i \subseteq A_i \times B_i$ so that the E_is are disjoint and cover the same set E of edges. Then

$$\mathrm{cc_w}(G) = \sum_{i=1}^{t}(|A_i| + |B_i|) = \sum_{i=1}^{t} \sum_{e \in E_i} \frac{|A_i| + |B_i|}{|E_i|} \ge \sum_{i=1}^{t} \sum_{e \in E_i} w(G) = w(G) \cdot |E|.$$

To prove the last claim, let $G = (V, E)$ be an extension of $H \subseteq U \times W$; hence, $E \cap (U \times W) = H$. If $A_i \times B_i$, $i = 1, \dots, t$ is a bipartite clique covering of G, then $(A_i \cap U) \times (B_i \cap W)$, $(B_i \cap U) \times (A_i \cap W)$, $i = 1, \dots, t$ is a bipartite clique covering of H. The number of bipartite cliques in this new covering is at most twice of that in the original covering, and the total number of vertices in the new covering does not increase at all. □

If we consider circuits with gates of *arbitrary* fanout, then $\mathrm{cc_w}(G)$ may be larger than $C^1(f_G)$, because the collection of boolean sums (ORs of variables) on the first level (before AND gates) may be not necessarily computed separately: one partial sum computed at some OR gate may be used many times.

Let $\mathcal{A} = \{A_1, \dots, A_m\}$ be a family of subsets of $\{1, \dots, n\}$. The *disjunctive complexity* of $\mathcal{A}$ is the minimum number of gates in a circuit over the basis $\{\vee\}$ needed to simultaneously compute all m boolean sums

$$\bigvee_{i \in A_1} x_i, \; \dots, \; \bigvee_{i \in A_m} x_i.$$

REMARK 12. For every subset $I \subseteq \{1, \dots, n\}$, the disjunctive complexity of the restriction $\mathcal{A}_I = \{A_1 \cap I, \dots, A_m \cap I\}$ of $\mathcal{A}$ onto I does not exceed that of the original family $\mathcal{A}$: having a circuit for $\mathcal{A}$ we can get a circuit for $\mathcal{A}_I$ just by setting to 0 all variables x_i with $i \notin I$.

LEMMA 13 (Pudlák–Rödl–Savický [16]). *Let $\mathcal{A} = \{A_1, \dots, A_m\}$ be a family of subsets of $\{1, \dots, n\}$. Then the disjunctive complexity of $\mathcal{A}$ does not exceed $n + k2^{m+1} - m - 2$. In particular, the disjunctive complexity of $\mathcal{A}$ does not exceed $kn + k2^{\lceil m/k \rceil + 1}$ for every $k \geq 1$.*

Hence, the disjunctive complexity of some families may be much smaller than $\sum_{i=1}^{m} |A_i|$. Still, the overlap of gates cannot be too large if the sums are "disjoint enough". A family is (h, k)-*disjoint* if no $h + 1$ of its members share more than k elements in common.

LEMMA 14 (Wegener [22]). *If $\mathcal{A} = \{A_1, \dots, A_m\}$ is (h, k)-disjoint then $\mathcal{A}$ has disjunctive complexity at least*

$$\frac{1}{kh} \sum_{i=1}^{m} |A_i| - \frac{m}{h}.$$

Proof. [Sketch] At least $|A_i| - 1$ gates are necessary for computation of the i-th sum and at least $|A_i|/k - 1$ of the functions computed at these gates are boolean sums of more than k summands. We only count these gates. Since the family is (h, k)-disjoint, each of these gates can be useful for at most h outputs. Hence, we need at least $\sum_{i=1}^{m} (|A_i|/k - 1)/h$ gates to compute all m sums. □

Amano and Maruoka [2] used this lemma to prove that $C^1(f_G) \geq |E|$ for any graph $G = (V, E)$ with no copies of $K_{2,2}$. The argument of [2] can be easily adopted to obtain a lower bound $C^1(f_G) \geq \Omega(|E|/t^3)$ for graphs with no copy of $K_{t,t}$. However, we need a similar result for graphs G which are *saturated extensions* of bipartite $n \times n$ graphs, and such graphs already have copies of $K_{t,t}$ with $t = n/2$. The following lemma works also for graphs with large bipartite cliques.

LEMMA 15. *Let $H \subseteq U \times W$ be a bipartite $n \times n$ graph with no copies of $K_{t,t}$ and let G be an extension of H. Then*

$$C^1(f_G) = \Omega\left(\frac{|H| - tn}{t^3}\right).$$

Moreover, if H is star-free then $C^1(f_G) = \Omega\left(|H|/t^3\right)$.

Proof. Take a minimal monotone single level circuit F representing H. Any circuit computing the quadratic function f_G of an extension of H must clearly represent H, implying that $C^1(f_G)$ must be at least the size, size(F), of F. Since the graph H has no copies of $K_{t,t}$, it can have at most $2(t-1)$ complete stars. By setting to 0 all variables corresponding to centers of these stars we obtain a single level circuit F_1 representing an induced star-free subgraph H_1 of H with $|H_1| \geq |H| - 2tn$ edges. Moreover, size$(F) \geq$ size(F_1).

The circuit F_1 has the form $\bigvee_{i=1}^m g_i \wedge h_i$ where m is the number of AND gates in F_1, and

$$g_i = \bigvee_{u \in S_i} x_u \qquad \text{and} \qquad h_i = \bigvee_{v \in T_i} x_v$$

are boolean sums computed at the inputs of the i-th AND gate with $S_i, T_i \subseteq U \cup W$. Our goal is to show that we need many OR gates to compute these sums. We cannot apply Lemma 14 directly to these sums because the corresponding families may be not disjoint enough. Still, we can use the absence of $K_{t,t}$ in H_1 to show that the restriction of these families onto U or onto W must contain a large enough (t,t)-disjoint subfamily.

First, observe that $S_i \cap T_i = \emptyset$ because the graph H_1 is star-free. Also, if for some i, both S_i and T_i would entirely lie in the same part of the bipartition, then we could just remove the i-th AND gate—the resulting circuit would still represent H_1. So, we may assume that this does not happen. Hence, H_1 is the union of bipartite cliques $(S_i \cap U) \times (T_i \cap W)$ and $(T_i \cap U) \times (S_i \cap W)$ for $i = 1, \ldots, m$. We may assume w.l.o.g. that the union H_1' of cliques

$$A_i \times B_i = (S_i \cap U) \times (T_i \cap W),$$

$i = 1, \ldots, m$ contains at least $|H_1'| \geq |H_1|/2$ edges of H_1 (if not, then take the remaining bipartite cliques).

Since H_1' has no copies of $K_{t,t}$, for every $i = 1, \ldots, m$, at least one of the sets A_i and B_i must have fewer than t elements. Hence, if we set $I = \{i : |A_i| < t\}$ then $|B_i| < t$ for all $i \notin I$. We may assume that the bipartite graph

$$H_2 = \bigcup_{i \in I} A_i \times B_i$$

contains at least $|H_2| \geq |H_1'|/2 \geq |H_1|/4$ edges of H_1 (if not, then let H_2 be the union of bipartite cliques $A_i \times B_i$ with $i \notin I$ and replace the roles of A_i's and B_i's). We now collect the boolean sums $h_i, i \in I$ computed in F_1 into a circuit F_2, by the following construction

$$F_2(X) = \bigvee_{u \in A} x_u \wedge \Big(\bigvee_{i \in I_u} h_i \Big)$$

where $A = \bigcup_{i \in I} A_i$ and $I_u = \{i \in I : u \in A_i\}$. For every $u \in A$ and $v \in W$, the circuit F_2 accepts the arc uv iff $v \in T_i \cap W = B_i$ for some $i \in I$ such that $u \in A_i$. Hence, F_2 represents the graph H_2. For every $u \in A$, the sub-circuit $\bigvee_{i \in I_u} h_i$ computes the boolean sum $\bigvee_{v \in T_u} x_v$ with $T_u = \bigcup\{T_i : i \in I_u\}$. Hence, size$(F_2)$ is at least the disjunctive complexity of the family $\mathcal{T} = \{T_u : u \in A\}$ which, in its turn, is at least the disjunctive complexity of the restriction $\mathcal{T}' = \{T_u \cap W : u \in A\}$ of $\mathcal{T}$ onto the set W. Moreover, we have that $|H_2| = \sum_{u \in A} |T_u \cap W|$ since $T_u \cap W$ is the set of all neighbors of u in H_2. Since H_2 has no copies of $K_{t,t}$ and the circuit H_2 must reject all non-edges of H_2, the family $\mathcal{T}'$ must be (t,t)-disjoint. By Lemma 14,

$$\text{size}(F_2) \geq \frac{1}{t^2} \sum_{u \in A} |T_u'| - \frac{|A|}{t} = \frac{|H_2|}{t^2} - \frac{|A|}{t}.$$

It remains to show that the size of F_2 is not much larger than the size of the original circuit F. Since all boolean sums h_i with $i \in I$ are already computed in F_1, we need at most $\sum_{u \in A} |I_u| = \sum_{i \in I} |A_i| \leq t \cdot |I|$ new gates to compute all functions $x_u \wedge \big(\bigvee_{i \in I_u} h_i\big)$ with $u \in A$. To compute the disjunction of these functions we need at most $|A| \leq \sum_{i \in I} |A_i| \leq t \cdot |I|$ additional OR gates. Hence, size$(F_2) \leq$ size$(F_1) + 2t \cdot |I| \leq 3t \cdot$ size(F_1), implying that

$$\text{size}(F_1) \geq \frac{1}{3t} \cdot \text{size}(F_2) \geq \frac{|H_2|}{3t^3} - \frac{|A|}{3t^2} \geq \frac{|H_2|}{3t^3} - \text{size}(F_1).$$

Since $|H_2| \geq |H_1|/4 \geq |H|/4 - tn/2$ and size$(F) \geq$ size(F_1), the desired lower bound size$(F) = \Omega(|H|/t^3 - n/t^2)$ follows. $\square$

Now we turn to the actual proof of our main results. We first prove Theorem 3 (multiplicative complexity of circuits) and Theorem 6 (combinational complexity of formulas); the proofs here are relatively simple. We then turn to the proof of Theorem 4 (combinational complexity of circuits).

4. Multiplicative complexity: proof of Theorem 3. Let G be a saturated extension of an n to n matching $M_n \subseteq U \times W$. Then, by Lemma 11, $C^1_{\mathrm{mult}}(f_G) = \mathrm{cc}(G) \geq \mathrm{cc}(M_n)/2 = n/2$. On the other hand, as shown in [9], M_n can be represented by a monotone CNF with $O(\log n)$ clauses: let $r = 2\log n$ and associate with each vertex u_i on the left side its *own* $r/2$-element subset A_i of $\{1,\ldots,r\}$, and assign to the unique matched vertex v_i on the right side the complement $B_i = \overline{A_i}$ of this subset. It is clear that then $A_i \cap B_j = \emptyset$ iff $i = j$. By Fact 8, $\mathrm{cnf}(M_n) = \mathrm{int}(M_n) \leq r = 2\log n$. Together with Lemma 10, this implies that $C_{\mathrm{mult}}(f_G) = O(\log n)$. Hence, $\mathrm{Gap}_{\mathrm{mult}}(G) = \Omega(n/\log n)$. $\Box$

5. Formulas: proof of Theorem 6. Let G be the saturated extension of the bipartite Kneser $n \times n$ graph $K \subseteq U \times V$. Recall that in this case U and W consist of all $n = 2^r$ subsets u of $\{1,\ldots,r\}$, and $uv \in K$ iff $u \cap v = \emptyset$. Since $\log_2 3 > 1.58$, the graph K has

$$|K| = \sum_{u\in U} d(u) = \sum_{u\in U} 2^{r-|u|} = \sum_{i=0}^{r} \binom{r}{i} 2^{r-i} = 3^r \geq n^{3/2+c}$$

edges with $c \geq 0.08$. Moreover, the graph K can contain a complete bipartite $a \times b$ subgraph $\emptyset \neq A \times B \subseteq K$ only if $a \leq 2^k$ and $b \leq 2^{r-k}$ for some $0 \leq k \leq r$, because then it must hold $\left(\bigcup_{u\in A} u\right) \cap \left(\bigcup_{v\in B} v\right) = \emptyset$. Since $a \leq a'$ and $b \leq b'$ imply $(a+b)/ab \geq (a'+b')/a'b'$, we have

$$w(K) \geq \frac{2^k + 2^{r-k}}{2^r} \geq 2^{-r/2} \geq n^{-1/2}.$$

By Lemma 11, $L^1(f_G) = \mathrm{cc_w}(G) \geq \mathrm{cc_w}(K) \geq w(K) \cdot |K| \geq n^{1+c}$.

On the other hand, by its definition, the graph K has intersection representation of size r and, by Fact 8, can be represented by a monotone CNF with $\mathrm{int}(K) \leq r = \log(n+1)$ clauses. Hence K, can be represented by a monotone formula with $O(n\log n)$ fanin-2 AND and OR gates. Together with Lemma 10, this implies that $L(f_G) = O(n\log n)$. Hence, $\mathrm{Gap}_{\mathrm{form}}(G) = L^1(f_G)/L(f_G) = \Omega(n^c/\log n)$. $\Box$

6. Circuits: proof of Theorem 4. Let $\mathbb{F} = \mathrm{GF}(2)$ and r be a sufficiently large even integer. With every subset $S \subseteq \mathbb{F}^r$ we associate a bipartite graph $H_S \subseteq S \times S$ such that two vertices u and v are adjacent if and only if $u \cdot v = 1$, where $u \cdot v$ is the scalar product over $\mathbb{F}$. We will need the following Ramsey-type property of such graphs.

LEMMA 16 (Pudlák–Rödl [15]). *Suppose every vector space $V \subseteq \mathbb{F}^r$ of dimension $\lfloor (r+1)/2 \rfloor$ intersects S in less than t elements. Then neither H_S nor the bipartite complement $\overline{H}_S$ contains $K_{t,t}$.*

Proof. [Sketch] The proof is based on the observation that any copy of $K_{t,t}$ in G_S would give us a pair of subsets X and Y of S of size t such that $x \cdot y = 1$ for all $x \in X$

and $y \in Y$. Looking the vectors in X as the rows of the coefficient matrix and vectors in Y as unknowns, we obtain that the sum $\dim(X') + \dim(Y')$ of the dimensions of vector spaces X' and Y', spanned by X and by Y, cannot exceed $r + 1$. Hence, at least one of these dimensions is at most $(r + 1)/2$, implying that either $|X' \cap S| < t$ or $|Y' \cap S| < t$. However, this is impossible because both X' and Y' contain subsets X and Y of S of size t. $\square$

Together with a simple probabilistic argument, this lemma yields

LEMMA 17. *There exists a subset $S \subseteq \mathbb{F}^r$ of size $|S| = 2^{r/2}$ such that neither H_S nor the bipartite complement $\overline{H}_S$ contains a copy of $K_{t,t}$ with $t = \omega(r)$.*

Proof. Let $N = 2^r$, and let $\mathbf{S} \subseteq \mathbb{F}^r$ be a random subset where each vector $u \in \mathbb{F}^r$ is included in $\mathbf{S}$ independently with probability $p = 2^{1-r/2} = 2/\sqrt{N}$. By Chernoff's inequality, $|\mathbf{S}| \geq pN/2 = 2^{r/2}$ with probability at least $1 - e^{-\Omega(pN)} = 1 - o(1)$.

Let now $V \subseteq \mathbb{F}^r$ be a subspace of $\mathbb{F}^r$ of dimension $\lfloor (r + 1)/2 \rfloor = r/2$ (remember that r is even). Then $|V| = 2^{r/2} = \sqrt{N}$ and we may expect $\mu = p|V| = 2$ elements in $|\mathbf{S} \cap V|$. By Chernoff's inequality (see, e.g. [14], Theorem 4.1),

$$\Pr\{|\mathbf{S} \cap V| \geq (1 + \lambda)\mu\} \leq e^{-\min(\lambda^2,\lambda)\mu/3}$$

holds for any $\lambda \geq 0$. The number of vector spaces in $\mathbb{F}^r$ of dimension $r/2$ does not exceed $\binom{r}{r/2} \leq 2^r/\sqrt{r}$. Hence, we can take $\lambda = 3r/2$ and conclude that the set $\mathbf{S}$ intersects some $r/2$-dimensional vector space V in $t = (1 + \lambda)\mu = 3r + 2$ or more elements with probability at most $e^{r-(\ln r)/2-r} \leq r^{-1/2} = o(1)$. Hence, with probability $1 - o(1)$ the set $\mathbf{S}$ has size at least $2^{r/2}$ and $|\mathbf{S} \cap V| < t$ for every $r/2$-dimensional vector space V. Fix such a set S and take an arbitrary its subset $S' \subseteq S$ of size $|S'| = 2^{r/2}$. By Lemma 16, neither $H_{S'}$ nor $\overline{H}_{S'}$ contains a copy of $K_{t,t}$ with t larger than $O(r)$. $\square$

Now we turn to the actual proof of Theorem 4.

Let $S \subseteq \mathbb{F}^r$ be a subset of size $|S| = n = 2^{r/2}$ guaranteed by Lemma 17. We may assume that $u \cdot v = 1$ holds for at least half of the pairs in S (otherwise take the bipartite complement of H_S). Hence, $H' = H_S$ is a bipartite $n \times n$ graph with $n = |S|$ vertices in each part and with $|H'| \geq |S|^2/2 = \Omega(n^2)$ edges. Moreover, this graph can contain a copy of $K_{t,t}$ only if $t = O(r) = O(\log n)$.

Let G be the saturated extension of H'. By Lemma 15,

$$C^1(f_G) = \Omega\left(\frac{|H'|}{t^3} - \frac{n}{t^2}\right) = \Omega\left(\frac{n^2}{\log^3 n}\right).$$

To get an upper bound on $C(f_G)$, let us identify each vector $w \in S$ with the set of 1-coordinates of w. Hence, two vertices u and v are adjacent in H' iff $|u \cap v|$ is odd. It is not difficult to see that (for even r) the graph H' can be represented by a depth-2 formula

$$F(X) = \bigoplus_{i=1}^{r} \bigvee_{w \in S_i} x_w$$

with $S_i = \{w \in S : i \notin w\}$. Indeed, the i-th clause $\bigvee_{w \in S_i} x_w$ accepts an arc $uv \in S \times S$ iff $u \in S_i$ or $v \in S_i$ iff $i \notin u \cap v$. Hence, the formula F accepts uv iff uv is accepted by an odd number of clauses iff $|\{i : i \notin u \cap v\}| = r - |u \cap v|$ is odd iff $|u \cap v|$ is odd iff $uv \in H'$.

By Lemma 13, all $r = 2\log n$ boolean sums in the formula $F(X)$ above can be computed using $O(n)$ fanin-2 OR gates. Hence, the graph H' can be represented by a circuit over the basis $\{\vee, \wedge, \neg\}$ consisting of $O(n)$ fanin-2 gates and, by Lemma 9, can be represented by a monotone circuit of size $O(n)$. Since G is the saturated extension of H', Lemma 10 implies that $C(f_G) = O(n)$. Hence, $\mathrm{Gap}(G) = C^1(f_G)/C(f_G) = \Omega\left(n/\log^3 n\right)$. ∎

7. Unbounded fanin circuits: proof of Theorem 7. First we show that the presence of unbounded fanin gates may *exponentially* increase the power even of single level circuits. For a graph G, let $C_*(G)$ (resp., $C_*(G)$) be the minimum size of a monotone unbounded fanin circuit (resp., single level circuit) representing G.

FACT 18. *Let H be a bipartite $n \times n$ graph of maximal degree $d < n$, and G be its saturated extension. Then $C^1_*(f_G) \le C^1_*(H) + O(\log n)$ and $C^1_*(H) \le \mathrm{cnf}(H) = \mathrm{int}(H) = O(d^2 \log n)$.*

Proof. The first claim is a direct consequence of Lemma 10. The inequality $C^1_*(H) \le \mathrm{cnf}(H)$ is trivial, and the equality $\mathrm{cnf}(H) = \mathrm{int}(H)$ is Fact 8. The upper bound $\mathrm{int}(H) = O(d^2 \log n)$ follows from the upper bound $O(d^2 \log n)$ on the clique covering number of the complement of H, due to Alon [1], and an observation that this number coincides with $\mathrm{int}(G)$ made by Erdős, Goodman and Pósa in [7]. ∎

This fact implies that $C^1_*(f_G)$ is exponentially smaller than $C^1(f_G)$ for a large class of graphs: if, say, G is the saturated extension of a constant degree bipartite $n \times n$ graph H without isolated vertices, then $C^1(f_G) \ge \Omega(|H|) = \Omega(n)$ (by Lemma 15) but $C^1_*(f_G) = O(\log n)$ (by Fact 18). Hence, the presence of unbounded fanin gates can indeed exponentially increase the power of single level circuits. This also shows that, for some graphs G, $C^1_*(f_G)$ may be exponentially smaller than the number of variables the function f_G depends on.

Theorem 7 is a direct consequence of the following general lower bound on the size of single level circuits with unbounded fanin gates.

Since the quadratic function of any bipartite complete graph $A \times B$ can be computed by a monotone CNF $\left(\bigvee_{u \in A} x_u\right) \wedge \left(\bigvee_{v \in B} x_v\right)$ of length 2, we have that $L^1_*(f_G) \le 3 \cdot \mathrm{cc}(G) + 1$ holds for all graphs G. On the other hand, if G has maximum degree d then $|E|/d^2$ is a trivial lower bound on $\mathrm{cc}(G)$. It turns out that this is also a lower bound on $L^1_*(f_G)$.

THEOREM 19. *If $G = (V, E)$ is a star-free graph of maximum degree d, then $L^1_*(f_G) \ge |E|/d^2$ and $C^1_*(f_G) \ge \sqrt{|E|}/d$.*

For the proof we need the following lemma. Let $\mathrm{cnf}(f_G)$ denote the minimum length of (i.e. the number of clauses in) a monotone CNF computing f_G.

LEMMA 20. *If H is a star-free graph with M edges and maximum degree d, then* $\mathrm{cnf}(f_H) \geq M/d^2$.

Proof. Let F be a monotone CNF of length $t = \mathrm{cnf}(f_H)$ computing f_H. Since H has no complete stars, this CNF must contain at least two clauses. Take any of these clauses $C = \bigvee_{u \in S} x_u$ and consider the shrinked CNF $F' = F \setminus \{C\}$. Since C must accept all edges of H, each of these edges must have at least one endpoint in S. But any one vertex in S can be an endpoint of at most d edges, implying that $|S| \geq M/d$.

Since F is a shortest CNF computing f_H, the shrinked CNF F' must make an error, i.e. it must (wrongly) accept some independent set of H. That is, there must be an independent set I such that every clause of F' contains a variable x_v with $v \in I$. Since F' has only $t-1$ clauses, we may assume that $|I| \leq t-1$. This error must be corrected by the clause C, implying that every vertex $u \in S$ must be adjacent (in H) with at least one vertex in I, for otherwise F would wrongly accept the independent set $I \cup \{u\}$ of H. Hence, at least one vertex $v \in I$ must have at least $|S|/|I| \geq M/td$ neighbors in H. Since the degree of v cannot exceed d, the desired lower bound $t \geq M/d^2$ follows. □

Proof of Theorem 19. We first consider the case of formulas. Let F be a smallest monotone Σ_3 formula computing f_G. This formula is an OR $F = F_1 \vee \cdots \vee F_s$ of monotone CNFs, and $\mathrm{size}(F) \geq \sum_{i=1}^{s} r_i$ where r_i is the length of the i-th CNF F_i. The CNFs F_i, $i = 1, \ldots, s$ compute quadratic functions of subgraphs $G_i = (V, E_i)$ of G such that $E_1 \cup \cdots \cup E_s = E$. Since each of these subgraphs is star-free and has maximum degree at most d, Lemma 20 yields $\mathrm{size}(F) \geq \sum_{i=1}^{s} r_i \geq \sum_{i=1}^{s} |E_i|/d^2 \geq |E|/d^2$.

If F is not a formula (some OR gates on the bottom level have fanout larger than 1), then we still have that $\mathrm{size}(F) \geq t = \max\{s, r_1, \ldots, r_s\}$. To get the desired lower bound $t \geq \sqrt{|E|}/d$, take a CNF F_i whose graph $G_i = (V, E_i)$ contains the maximal number of edges. By Lemma 20, F_i has length $r_i \geq |E_i|/d^2 \geq |E|/sd^2$. Since both r_i and s do not exceed t, this yields $t^2 \geq |E|/d^2$, and the desired lower bound $t \geq \sqrt{|E|}/d$ on the number of gates in F follows. □

Since $L_*(f_G) \leq |E| + 1$ for every graph $G = (V, E)$, Theorem 19 implies that $\mathrm{Gap}^*_{\mathrm{form}}(G) = O(d^2)$ for every star-free graph G of maximum degree d. This means that, also in the case of unbounded fanin formulas, the single level conjecture holds for many graphs (e.g. for graphs of constant degree)—for such graphs we have $\mathrm{Gap}^*_{\mathrm{form}}(G) = O(1)$. However, if this would hold for *all* graphs, then we would have a consequence (stated in Theorem 7) that for *every* star-free graph $G = (V, E)$ of maximum degree d, the quadratic function f_G cannot be computed by a circuit using fewer than $\Omega(\sqrt{|E|}/d)$ unbounded fanin AND and OR gates and cannot be computed by a formula using fewer than $\Omega(|E|/d^2)$ such gates. This gives an (indirect) indication that the single level conjecture should fail also for unbounded fanin circuits or formulas. Still, we cannot exclude this (rather unlikely) general lower bound on the size of monotone unbounded fanin circuits for quadratic functions.

PROBLEM 21. *Exhibit a star-free n-vertex graph G with $L_*(f_G) = o\left(|E|/d^2\right)$ or $C_*(f_G) = o(\sqrt{|E|}/d)$ (or prove a mere existence of such graphs).*

8. Open problems. Besides the problem above, there are many other interesting problems concerning the circuit complexity of graphs and their quadratic functions. Here we describe some of them.

In Theorem 4 we show a mere existence of a graph G with $\mathrm{Gap}(G) = \Omega(n/\log^4 n)$.

PROBLEM 22. *Exhibit an explicit n-vertex graph G with* $\mathrm{Gap}(G) = n^{\Omega(1)}$.

For a graph G, let $C(G)$ be the minimum size of a *monotone* circuit representing G. This measure is interesting because it is related to the *non-monotone* complexity of boolean functions. As we already mentioned in the introduction, lower bounds on the size of *monotone* circuits representing a graph G yield lower bounds on the *non-monotone* circuit size of their characteristic functions. In particular, a lower bound $C(G) \geq 12n + n^{\varepsilon}$ for an explicit bipartite $n \times n$ graph G would yield a lower bound $2^{\varepsilon m}$ on the *non-monotone* circuit size of an explicit boolean function in $2m$ variables (see [16, 9]). It is therefore not surprising that the measure $C(G)$ is not easy to deal with. Note that, by easy counting, $C(G) = \Omega(n^2/\log n)$ for almost all bipartite $n \times n$ graphs. The problem, however, is the *explicitness*: we want a lower bound for explicitly constructed graphs.

The monotone complexity of quadratic functions f_G is a more tractable measure (cf. Theorem 5). Even more, for some graphs (like saturated ones), f_G is the only monotone boolean function representing G. Hence, proving lower bounds on $C(f_G)$ for saturated graphs is of particular interest. For this, it would be interesting to better understand the connection between these two measures, $C(G)$ and $C(f_G)$. Since f_G represents G, we have that $C(G) \leq C(f_G)$ holds for any graph G and, by Fact 2, $C(G) = C(f_G)$ for saturated graphs.

PROBLEM 23. *For what graphs G, besides the saturated extensions, do we have that $C(G) = \Omega(C(f_G))$?*

Note, however, that for some graphs G, it may be much easier to represent G then to compute f_G. For example, if M is an n to n matching, then $\mathrm{cnf}(M) = O(\log n)$ (see § 4) but, by Theorem 19, every monotone Σ_3 circuit for f_M requires size $\sqrt{n}$. Hence, $C_*^1(M)$ is exponentially smaller than $C_*^1(f_M)$.

Explicit constructions of bipartite $n \times n$ graphs G without a copy of $K_{2,2}$ and with $\Omega(n^{3/2})$ edges are well known (see, e.g. [10]). For these graphs Theorem 5 yields a lower bound $L(f_G) = \Omega(n^{3/2})$. However, proving lower bounds on the monotone size of quadratic functions in the case of *circuits* (arbitrary fanout) is a more difficult task. Razborov's argument [17] works well only if *both* minterms and maxterms are short enough. In the case of quadratic functions this is not the case: here minterms have length 2 but maxterms may be very large. Hence, for functions of the form f_G we need some new arguments.

Problem 24. *Prove $C(f_G) = \Omega(n^{1+\varepsilon})$ for an explicit n-vertex graph G.*

It is an interesting open question on whether the single level conjecture holds in the context of *graph representation.* If we allow only fanin-2 gates, then the conjecture is false: in § 6 we have shown that there are bipartite $n \times n$ graphs H such that $C(H) = O(n)$ and $C^1(H) = \Omega(n^{2-\varepsilon})$. But what happens if we allow gates of *unbounded* fanin?

Problem 25 (Pudlák–Rödl–Savický [16]). *Prove that $C_*(G)$ may be much smaller than $C_*^1(G)$.*

Unlike for quadratic functions f_G (cf. Theorem 19), proving non-trivial lower bounds on $C_*^1(G)$ seems to be a much more difficult task.

Problem 26 ([9]). *Exhibit an n-vertex graph G with $C_*^1(G) = n^{\Omega(1)}$.*

By a well-known result of Valiant [20] this would yield a super-linear lower bound for non-monotone log-depth circuits (see [9] for details), thus solving an old problem in circuit complexity.

Acknowledgment. I would like to thank Ingo Wegener for useful comments.

REFERENCES

[1] Alon, N. (1986): Covering graphs by the minimum number of equivalence relations, *Combinatorica* **6**, 201–206.

[2] Amano, M. and Maruoka, A. (2004): On the monotone circuit complexity of quadratic Boolean functions. In: *Proc. of 5-th Int. Symp. on Algorithms and Computation*, ISAAC 2004 (Hong Kong, China, December 20-22), Lect. Notes in Comput. Sci., vol. 3341, 28–40.

[3] Bloniarz, P. A. (1977): *The complexity of monotone boolean functions and an algorithm for finding shortest paths in a graph*, PhD Thesis, MIT.

[4] Bublitz, S. (1986): Decomposition of graphs and monotone size of homogeneous functions, *Acta Informatica* **23**, 689–696.

[5] Chung, F. R. K. (1980): On the covering of graphs, *Discrete Math.* **30**, 89–93.

[6] Chung, F. R. K., Erdős, P. and Spencer, J.(1983): On the decomposition of graphs into complete bipartite subgraphs. In: *Studies in pure mathematics*, Mem. of P. Turán, pp. 95–101.

[7] Erdős, P., Goodman, A. W. and Pósa, L. (1966): The representation of a graph by set intersections, *Can. J. Math.* **18**, 106–112.

[8] Dunne, P. (1984): *Techniques for the analysis of Boolean networks*, PhD Thesis, Univ. Warwick.

[9] Jukna, S. (2004): On graph complexity, *Electronic Colloq. on Comp. Complexity*, Report Nr. 5, 18 pp.

[10] Kövári, P., Sós, V.T. and Túran, P. (1954): On a problem of Zarankiewicz, *Colloq. Math* **3**, 50–57.

[11] Krichevski, R.E. (1964): Complexity of contact circuits realizing a function of logical algebra, *Sov. Phys., Dokl.* **8**, 770–772; translation from *Dokl. Akad. Nauk SSSR* **151** (1963), 803–806.

[12] Lenz, K. and Wegener, I. (1991): The conjunctive complexity of quadratic boolean functions, *Theor. Comput. Sci.* **81**, 257–268.

[13] Mirwald, R. and Schnorr, C.P. (1992): The multiplicative complexity of quadratic boolean forms, *Theor. Comput. Sci.* **102**:2, 307–328 (1992).

[14] Motwani,R. and Raghavan, P. (1995): *Randomized Algorithms*, Cambridge University Press, New York.

[15] Pudlák and P., Rödl, V. (2004): Pseudorandom sets and explicit constructions of Ramsey graphs. In: *Quaderni di Matematica* (to appear)

[16] Pudlák, P., Rödl, V. and Savický, P. (1988): Graph complexity, *Acta Informatica* **25**, 515–535.

[17] Razborov, A. A. (1985): Lower bounds on the monotone complexity of some Boolean functions, *Doklady Akademii Nauk SSSR* **281**:4, 798–801. English translation in: *Soviet Mathematics Doklady,* **31** (1985), 354–357.

[18] Rödl, V. and Ruciński, A. (1997): Bipartite coverings of graphs, *Combinatorics, Probability* & *Computing* **6**, 349–352.

[19] Tuza, Z. (1984): Covering of graphs by complete bipartite subgraphs, complexity of 0-1 matrices, *Combinatorica* **4**, 111–116.

[20] Valiant, L. (1977): Graph-theoretic methods in low-level complexity. In: *Proc. of* 6-*th Conf. on Mathematical Foundations of Computer Science*, Springer Lect. Notes in Comput. Sci., vol. 53, 162–176.

[21] Wegener, I. (1980): A new lower bound on the monotone network complexity of Boolean sums, *Acta Informatica* **15**, 147–152.

[22] Wegener, I. (1986): More on the complexity of slice functions, *Theor. Comput. Sci.* **43**, 201–211.

[23] Wegener, I. (1987): *The Complexity of Boolean Functions*, Wiley-Teubner.

ECCC ISSN 1433-8092
http://www.eccc.uni-trier.de/eccc
ftp://ftp.eccc.uni-trier.de/pub/eccc
ftpmail@ftp.eccc.uni-trier.de, subject 'help eccc'