Electronic Colloquium on Computational Complexity, Revision 2 of Report No. 21 (2005)

Disproving the single level conjecture *

Stasys Jukna †

Abstract

We consider the size of monotone circuits for quadratic boolean functions, that is, disjunctions of length-2 monomials. Our motivation is that a good (linear in the number of variables) lower bound on the monotone circuit size for a certain type of quadratic function would imply a good (even exponential) lower bound on the general non-monotone circuit size.

To get more insight into the structure of monotone circuits for quadratic functions, we consider the so-called single level conjecture posed explicitely around 1990. The conjecture claims that monotone single level circuits, that is, 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 whose monotone circuits have linear size whereas their monotone single level circuits require almost quadratic size.

1 Introduction

A quadratic boolean function is a monotone boolean function whose all prime implicants have length two. There is an obvious correspondence between such functions and graphs: every graph $G = (V, E)$ defines a natural quadratic function

$$f_G(X) = \bigvee_{uv \in E} x_u x_v, \tag{1}$$

and every quadratic function defines a unique graph. We consider the complexity of computing such functions by monotone circuits, that is, by 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. Note that

*Research supported by a DFG grant SCHN 503/2-2.

†Institute of Mathematics and Informatics, Akademijos 4, LT-08663 Vilnius, Lithuania. Current address: Universität Frankfurt, Institut für Informatik Robert-Mayer-Str. 11-15, D-60054 Frankfurt, Germany. Email: jukna@thi.informatik.uni-frankfurt.de

ISSN 1433-8092

every quadratic boolean function f_G in n variables can be computed by a trivial monotone single-level circuit with at most $n-1$ AND gates using the form

$$\bigvee_{u\in S} x_u \wedge \left(\bigvee_{v:uv\in E} x_v\right) \tag{2}$$

where $S \subseteq V$ is an arbitrary vertex cover of G, that is, a set of vertices such that every edge of G is incident with a vertex in S.

Single Level Conjecture: *For quadratic functions single level circuits are almost as powerful as unrestricted ones.*

Here "almost" means "up to a constant factor." This conjecture—first explicitly framed as the "single level conjecture" by Lenz and Wegener in [14]—was considered by several authors, [12, 4, 5, 17, 14, 2] among others. That the conjecture holds for *almost all* quadratic functions was shown by Bloniarz [4] more than twenty five years ago and, so far, no (even constant) gap between the size of general and single level circuits for quadratic functions was known.

In this paper we disprove the single level conjecture in a strong sense: there are quadratic functions in n variables whose monotone circuits have linear size whereas their monotone single level circuits require size $\Omega(n^2/\log^3 n)$. A similar gap is also shown for boolean formulas. We also discuss the single level conjecture in the case of monotone circuits with unbounded fanin gates.

Why should we care about monotone circuits for quadratic functions, when we already can prove high (even exponential) lower bounds for monotone circuits? There are several reasons for this.

1. Any explicit n-vertex graph G, that cannot be represented (in a sense described later in § 3) by a monotone circuit using fewer than cn gates for a sufficiently large constant $c > 0$, would give us an explicit *exponential* lower bound for general (non-monotone) circuits. Let us briefly sketch how this happens. 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 $f(uv) = 1$ if and only if $uv \in G$. Suppose now that we have a *non-monotone* circuit $F(y_1, \ldots, y_{2m})$ computing f whose inputs are variables y_i and their negations $\overline{y}_i$; the rest of the circuit is monotone (consists of AND and OR gates). Then, according to the so-called "magnification lemma" [10], it is possible to replace its $4m = 4\log n$ input literals (both positive and negative) by appropriate boolean sums (ORs) of variables in $X = \{x_v : v \in U \cup W\}$ so that the resulting *monotone* circuit $F_+(X)$ in $|X| = 2n$ variables represents G. It can be shown (see [21] or Lemma 3.6 below) that all these $4\log n$ boolean sums can be simultaneously computed by a monotone circuit of size cn for a constant c. Therefore, the size of F cannot be much smaller than that of F_+: $\mathrm{size}(F) \geq \mathrm{size}(F_+) - cn$. Hence, a lower bound $cn + n^\epsilon$ on the size of *monotone* circuits representing G would yield a lower bound $n^\epsilon = 2^{\epsilon m}$ on the *non-monotone* circuit size of an explicit boolean function f in $2m$ variables.

2. Better yet, for some graphs G, f_G is the *only* monotone boolean function representing G. Such are, in particular, complements of triangle-free graphs (see Observation 3.5

below). Hence, one could obtain large (even exponential) lower bounds for general non-monotone circuits by proving a good (but only linear) lower bound on the monotone circuit size of such quadratic functions.

3. Unlike boolean functions, graphs have been studied for a long time, and explicit constructions of graphs with very special properties are already known. It is therefore a hope to design a lower bound proof that is highly specialized for some particular graph or some small class of graphs. This could (probably) lead to a lower bound proof which will not fulfill the "largeness" condition in the notion of "natural proofs" [23].

4. When applied to quadratic functions, known lower bound arguments for monotone circuits—Razborov's method of approximations [22] and its modifications—cannot yield lower bounds larger than n. The reason for this is that these arguments are lower bounding the *minimum* of AND gates and that of OR gates needed to compute the function, and (as we already noted above) every quadratic boolean function f_G in n variables can be computed by a trivial monotone single-level circuit with at most n AND gates.

We therefore need entirely new lower bound arguments for monotone circuits computing quadratic functions. For this, it is important to better understand the *structure* of such circuits. And the (long studied) single level conjecture seems to be a good starting point in this direction.

2 Results

Let us first introduce some notation. By the *size* of a circuit we will always mean the number of gates in it. For a monotone boolean function f, let $C(f)$ denote the minimum number of gates and $C_{\&}(f)$ the minimum number of AND gates in a monotone circuit computing f. Let also $C^1(f)$ and $C^1_{\&}(f)$ denote the single level counterparts of these measures. Further, let $L(f)$ and $L^1(f)$ denote the minimum length of a monotone (resp., of a monotone single level) formula computing f. Recall that a *formula* is a circuit where all gates have fanout 1, i.e. the underlying graph is a tree; the *length* of the formula is the number of leaves of this tree.

In Table 1 we summarize known upper and lower bounds on the *maximum* possible complexity of quadratic functions f_G over all n-vertex graphs; the upper bounds here hold for all graphs and the lower bounds for almost all graphs.

Table 1: Known bounds on the maximum complexity of quadratic functions

Upper bounds	**Lower bounds**
$C^1(n) = O(n^2/\log n)$ ([4])	$C(n) = \Omega(n^2/\log n)$ ([4])
$L^1(n) = O(n^2/\log n)$ ([26, 5, 21])	$L(n) = \Omega(n^2/\log n)$
$C^1_{\&}(n) \leq n - \lfloor \log n \rfloor + 1$ ([26, 14])	$C^1_{\&}(n) \geq n - c \log n$ ([24])
	$C_{\&}(n) = \Omega(n/\log n)$ ([14])
	$C_{\&}(n) = \Omega(n)$ ([2])

In this paper we are interested in the corresponding gaps between general and single level complexities for *individual* graphs:

1. circuit gap $\mathrm{Gap}(G) = C^1(f_G)/C(f_G)$;
2. multiplicative gap $\mathrm{Gap}_{\mathrm{mult}}(G) = C^1_{\&}(f_G)/C_{\&}(f_G)$;
3. formula gap $\mathrm{Gap}_{\mathrm{form}}(G) = L^1(f_G)/L(f_G)$.

Note that the single level conjecture claims that $\mathrm{Gap}(G) = O(1)$ for all graphs G. Table 1 shows that, for almost all graphs, the conjecture is indeed true.

An even stronger support for the single level conjecture was given by Mirwald and Schnorr [17]: if we consider circuits over the basis $\{\oplus, \wedge, 1\}$ computing (algebraic) quadratic forms $\sum_{uv\in E} x_u x_v$ 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. But the case of circuits over the basis $\{\vee, \wedge, 0, 1\}$ remained unclear.

In the case of formulas, Krichevski [12] has proved that $\mathrm{Gap}_{\mathrm{form}}(K_n) = 1$ for the complete graph K_n on n vertices, even if negation is allowed as an operation. A graph with $\mathrm{Gap}_{\mathrm{form}}(G) \geq 8/7$ was given by Bublitz [5]. In the case of multiplicative complexity, a graph with $\mathrm{Gap}_{\mathrm{mult}}(G) \geq 4/3$ was given by Lenz and Wegener [14]. Recently, this gap was substantially enlarged to $\mathrm{Gap}_{\mathrm{mult}}(G) = \Omega(n/\log n)$ by Amano and Maruoka in [2]; this was implicit also in [10]. Using a construction of Tarjan [25] (which, in its turn, was used by Tarjan for disproving that AND gates are powerless for computing boolean sums), Amano and Maruoka [2] have also shown the gap $\mathrm{Gap}(\mathcal{F}) \geq 29/28$ for circuits computing a *set* $\mathcal{F}$ of quadratic functions. However, even the existence of a *single* graph G with $\mathrm{Gap}(G) > 1$ was not known.

Our main result is the following.

Theorem 2.1. *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 2.1 are saturated extensions of Sylvester-type graphs, that is, of bipartite graphs whose vertices are particular vectors in $\mathrm{GF}(2)^r$, and where two vertices are adjacent iff their scalar product over GF(2) is 1. The *saturated extension* of a bipartite graph $H \subseteq U \times W$ is a (non-bipartite) graph $G = (V, E)$ with $V = U \cup W$ such that $E \cap (U \times W) = H$ and the induced subgraphs of G on U as well as on W are complete graphs. The reason to consider graphs of this special form lies in the simple fact (Lemma 3.8 below) that having a small circuit representing H we can construct a small circuit computing f_G.

To disprove the single level conjecture for formulas, we consider a bipartite version of graphs introduced by Lovász [15] in his famous proof of Kneser's conjecture [11]. 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 2.2. *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,* $\mathrm{Gap}_{\mathrm{form}}(G) = n^{\Omega(1)}$.

Next, we consider the single level conjecture for monotone *unbounded fanin* circuits and formulas. Note that in this case single level circuits are precisely the Σ_3 circuits: 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 monotone boolean function f, let $C_*(f)$ (resp., $L_*(f)$) be the minimum size of a monotone unbounded fanin circuit (resp., formula) computing f. Let also $C^1_*(f)$ and $L^1_*(f)$ denote the corresponding measures in a class of monotone Σ_3 circuits (i.e. the single level versions of these measures). Note that, also in the case of formulas, we now count the number of gates, not the number of leaves.

Single level circuits of unbounded fanin are interesting for at least two reasons.

1. The presence of unbounded fanin gates may *exponentially* increase the power of single level circuits: if, say, G is the saturated extension of an n to n matching, then $C^1(f_G) = \Omega(n)$ but $C^1_*(f_G) = O(\log n)$ (by Lemmas 3.8, 3.10 and 3.13 below).

2. By the reduction due to Valiant [27], a lower bound of the form $n^{\Omega(1)}$ on the size of a monotone Σ_3 formula representing an explicit n-vertex graph would give us a *super-linear* lower bound on non-monotone (fanin-2) circuits of logarithmic depth, and thus, would resolve an old and widely open problem in circuit complexity (see [10] for details).

The form (2) implies that $C^1_*(f_G) = O(n)$ for all n-vertex graphs. On the other hand, easy counting shows $C_*(f_G) = \Omega(n)$ for almost all n-vertex graphs: every gate in a circuit of size t can have at most 2^t possible sets of immediate predecessors, implying an upper bound $2^{O(t^2)}$ on the total number of such circuits. Hence, also in the case of unbounded fanin circuits, the single level conjecture holds for *almost all* quadratic functions. The following theorem gives a stronger result: the conjecture holds for *explicit* (and large) classes of quadratic functions.

Recall that a set $S \subseteq V$ is a *vertex cover* of $G = (V, E)$ if every edge of G is incident with a vertex in S. Let $\tau(G)$ denote the minimum cardinality of a vertex cover of G. Note that for every n-vertex graph $G = (V, E)$ of maximum degree d we have $|E|/d \leq \tau(G) \leq n - 1$. Let also $\mathrm{m}(G)$ denote the maximum possible number m such that G contains a matching with m edges as an *induced* subgraph. Representation (2) gives us the upper bound $C^1_*(f_G) \leq L^1_*(f_G) \leq 2\tau(G) + 1$. On the other hand, we have the following lower bounds.

Theorem 2.3. *For every graph G we have $C_*(f_G) \geq \mathrm{m}(G) + 1$. Moreover, $L^1_*(f_G) \geq \tau(G)/d$ and $C^1_*(f_G) \geq \sqrt{\tau(G)/d}$, where d is the maximum degree of G.*

Hence, if we consider circuits with *unbounded* fanin gates, then the single level conjecture is true for all n-vertex graphs containing an induced matching with $\Omega(n)$ edges: for all such graphs we have $C_*(f_G) = \Omega(n)$ and $C^1_*(f_G) = O(n)$.

In the case of multiplicative complexity (where we count only AND gates) we have the following gap.

Theorem 2.4 ([2]; implicit in [1, 10])**.** *If G is the saturated extension of an n to n matching, then $C_{\&}(f_G) = O(\log n)$ but $C^1_{\&}(f_G) = \Omega(n)$. Hence,* $\mathrm{Gap}_{\mathrm{mult}}(G) = \Omega(n/\log n)$.

This result was implicit in [10] (and even in [1], cf. Lemma 3.10 below) where it was shown that an n to n matching M (a bipartite $n \times n$ graph consisting of n vertex disjoint edges) can be represented by a monotone CNF with $O(\log n)$ clauses. The proof in this case is particularly simple, and we include it just for completeness. Amano and Maruoka [2] have used a somewhat different argument to show the same gap.

Table 2: Summary of results concerning the single level conjecture

	Known	**This paper**
circuits (all gates)	$\mathrm{Gap}(\mathcal{F}) \geq 29/28$ ([2]) (for a set of graphs; no known gap for a single graph)	$\mathrm{Gap}(G) = \Omega(n/\log^3 n)$ Sylvester-type graphs (main result)
formulas (all gates)	$\mathrm{Gap}_{\mathrm{form}}(K_n) = 1$ ([12]) $\mathrm{Gap}_{\mathrm{form}}(G) \geq 8/7$ ([5])	$\mathrm{Gap}_{\mathrm{form}}(G) = n^{\Omega(1)}$ Kneser-type graphs
circuits (AND gates)	no gap over GF(2) [17] $\mathrm{Gap}_{\mathrm{mult}}(G) \geq 4/3$ ([14]) $\mathrm{Gap}_{\mathrm{mult}}(G) = O(n/\log\log n)$ ([2])	$\mathrm{Gap}_{\mathrm{mult}}(G) = \Omega(n/\log n)$ perfect matchings (also in [2]; implicit in [1, 10])
unbounded fanin (all gates)		$C_*(f_G) \geq \mathrm{m}(G) + 1$ $\tau(G)/d \leq L^1_*(f_G) \leq 2\tau(G) + 1$ $C^1_*(f_G) \geq \sqrt{\tau(G)/d}$ $\forall$ G of maximal degree d

The rest of the paper is organized as follows. In the next section we collect some preliminary definitions and technical facts. We then use these facts to prove Theorems 2.1–2.4 in §§ 4–7. We conclude with several open problems.

3 Preliminaries

In this section we first recall from [10] the notion of graph representation, expose some properties of quadratic functions of saturated graphs and recall some results about boolean sums. We then prove some general (graph theoretic) bounds on the circuit complexity of quadratic functions.

We shall use standard graph theory notation. A set of vertices is *independent* if no two of its vertices are adjacent. A *non-edge* is a pair of non-adjacent vertices; if the graph is bipartite then a non-edge is a pair of non-adjacent vertices from different parts (color classes), that is, pairs of vertices in one color class are neither edges nor non-edges. A *subgraph* (or a *spanning subgraph*) of a graph is obtained by deleting its edges. An *induced subgraph* is obtained by deleting vertices (together with all the edges incident with them). The main difference between these two types of subgraphs is that every non-edge of an induced subgraph is also a non-edge of the original graph. A bipartite clique $K_{a,b}$ is a complete bipartite graph with color classes of size a and b.

3.1 Graph representation

Every graph $G = (V, E)$ gives us a set of boolean functions "representing" this graph in the following sense. We associate to each vertex v a boolean variable x_v, and consider boolean functions $f(X)$ with $X = \{x_v : v \in V\}$. Such a function accepts/rejects a subset of vertices $S \subseteq V$ if it accepts/rejects the incidence vector of S. We are interested in the behavior of such functions on edges and non-edges of G, viewed as 2-element sets of their endpoints.

Definition 3.1 **([10]).** A boolean function *represents* a given graph if it accepts all edges and rejects all non-edges.

Hence, $f(X)$ represents the graph G if for every input vector $a \in \{0,1\}^X$ with precisely two 1's in, say, positions u and v, $f(a) = 1$ if uv is an edge, and $f(a) = 0$ if uv is a non-edge of G. If uv is neither an edge nor a non-edge (in the bipartite case) or if a contains more or less than two 1's, then the value $f(a)$ may be arbitrary.

Note that the quadratic function f_G represents the graph G in a strong sense: for every subset $S \subseteq V$, $f_G(S) = 0$ if and only if S is an independent set of G. But, in general, there may be many other boolean functions representing the same graph, because they do not need to reject independent sets with more than two vertices. Hence, there are more chances to design a small circuit representing a given graph than to (directly) design a small circuit computing its quadratic function. We will use this possibility later to upper bound the circuit size of quadratic functions.

A *complete star* around a vertex u in a graph with n vertices is a set of $n-1$ edges sharing u as one of their endpoints. 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. The only property of star-free graphs we will use later is given by the following simple

Observation 3.2. *Any monotone boolean function representing a star-free graph must reject all its single vertices.*

This is true because $f(\{u\}) = 1$ together with the monotonicity of f implies that f must accept all edges of a complete star around u.

3.2 Saturated graphs

As noted above, besides the quadratic function f_G, there may be many other monotone boolean functions representing G—these functions may "wrongly" accept some independent sets of G of cardinality 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 way we come to the following

Definition 3.3. A graph G is *saturated* if it has no independent sets with more than two vertices, that is, if the complement of G is a triangle-free graph.

The first interesting property of quadratic functions of saturated graphs is that these functions belong to a fundamental class of so-called "slice functions" were negation is almost powerless (see, e.g., [29], §§ 6.13-6.14). Recall that a *k-slice function* is a monotone boolean function f such that $f(a) = 0$ for inputs a with less than k ones, and $f(a) = 1$ for inputs a with more than k ones, that is, $f = f \wedge T_k^n \vee T_{k+1}^n$.

Observation 3.4. *If G is a saturated graph, then f_G is a 2-slice function.*

Proof. Let $G = (V, E)$ be a saturated graph, and $S \subseteq V$. If $|S| < 2$ then $f_G(S) = 0$, by the definition of quadratic functions (they cannot have prime implicants shorter than 2). If $|S| > 2$ then S cannot be an independent set since G is saturated; hence, $f_G(S) = 1$. □

The next interesting property of saturated graphs is their unique function representation.

Observation 3.5. *If G is a saturated star-free graph, then f_G is the only monotone boolean function representing G.*

Proof. Let f be an arbitrary monotone boolean function representing G. We have to show that $f(S) = f_G(S)$ for all subsets $S \subseteq V$. If $f_G(S) = 1$ then S contains both endpoints of some edge. 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, and $|S| \leq 2$ since G is saturated. Hence, S is either a single vertex or a non-edge. In both cases we have that $f(S) = 0$ because f must reject all non-edges and, by Observation 3.2, must also reject all single vertices. □

3.3 Boolean sums

We shall also use the following two facts about the monotone complexity of boolean sums. The *disjunctive complexity* of a collection of boolean sums $\bigvee_{i \in S_1} x_i, \ldots, \bigvee_{i \in S_m} x_i$ (or of the corresponding family of sets $S_1, \ldots, S_m$) is the minimum size of a circuit consisting solely of fanin-2 OR gates and simultaneously computing all these m boolean sums.

Lemma 3.6 (Pudlák–Rödl–Savický [21]). *For every $m \geq k \geq 1$, the disjunctive complexity of any family of m subsets of $\{1, \ldots, n\}$ does not exceed*

$$kn + k2^{\lceil m/k \rceil + 1}.$$

In particular, any collection of $k \log n$ boolean sums in n variables can be simultaneously computed by a circuit consisting solely of at most $3kn$ fanin-2 OR gates.

By this lemma, boolean sums may not necessarily be computed separately: one partial sum computed at some OR gate may be used many times. Still, the overlap of gates cannot be too large if the sums are "disjoint enough". A family of sets is (h, k)-*disjoint* if no $h + 1$ of its members share more than k elements in common.

Lemma 3.7 (Wegener [28], Mehlhorn [16]). *Any (h, k)-disjoint family $S_1, \ldots, S_m$ has disjunctive complexity at least*

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

Proof sketch. At least $|S_i| - 1$ gates are necessary for computation of the i-th sum and at least $|S_i|/k - 1$ 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 $\sum_{i=1}^{m}(|S_i|/k - 1)/h$ gates to compute all m sums.

3.4 Upper bounds for general circuits

An *extension* of a bipartite graph $H \subseteq U \times W$ is a (non-bipartite) graph $G = (V, E)$ with $V = U \cup W$ such that $E \cap (U \times W) = H$. The *saturated extension* is an extension whose induced subgraphs on U as well as on W are complete graphs. That is, saturated extensions consist of two disjoint cliques with some edges between these cliques. A useful property of such graphs (besides that they are saturated) is that the complexity of *computing* f_G cannot be much larger than the complexity of *representing* H: to determine the value $f_G(S)$ it is enough to additionally test whether S has more than two elements.

By the *length* of a CNF we mean the number of clauses in it.

Lemma 3.8. *Let $H \subseteq U \times W$ be a bipartite $n \times n$ graph, G the saturated extension of H, and f a monotone boolean function representing H. Then $f_G = (f \wedge g) \vee h$ where g is a monotone CNF of length 2 and h is an OR of $O(\log n)$ monotone CNFs of length 2. Moreover, if H is star-free then $f_G = f \vee h$.*

Remark: Note that $C_{\&}(h) = O(\log n)$, $L(h) = O(n \log n)$ and $C(h) = O(n)$. The first two upper bounds are obvious. The third follows from Lemma 3.6.

Proof. Let $g = (\bigvee_{u \in U} x_u) \wedge (\bigvee_{w \in W} x_w)$ and $h = K_U \vee K_W$ where $K_U(S) = 1$ iff $|S \cap U| \geq 2$, that is, K_U is the quadratic function of a complete graph on U. Since the edges of a complete graph n-vertex graph can be covered by $m \leq \lceil \log n \rceil$ bipartite cliques, each of the functions K_U and K_W has the form

$$\bigvee_{i=1}^{m} \Big(\bigvee_{u \in A_i} x_u \Big) \wedge \Big(\bigvee_{v \in B_i} x_v \Big) \tag{3}$$

with $m \leq \lceil \log n \rceil$ and $A_i \cap B_i = \emptyset$ for all $i = 1, \ldots, m$. Hence, h can be computed by an OR of m monotone CNFs of length 2. 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 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. Hence, the function $(f \wedge g) \vee h$ rejects S.

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

Lemma 3.8 gives us a simple (but useful) tool to show that a quadratic function f_G of the saturated extension of a bipartite graph H can be *computed* by a small monotone circuit: it is enough to *represent* H by a small circuit. To achieve this last goal, it is often enough to show that H has small "intersection representation."

Say that a graph G admits an *intersection representation of size* r if it is possible to associate with every vertex u 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 a non-edge of G. Let $\mathrm{int}(G)$ denote the smallest r for which G admits such a representation.

Let $\mathrm{cnf}(G)$ denote the minimum length of a monotone CNF representing the graph G, and let $\mathrm{cov}(G)$ denote the minimum number of independent sets of G covering all non-edges of G.

Lemma 3.9 ([9, 10])**.** *For every graph G, $\mathrm{cnf}(G) = \mathrm{int}(G) = \mathrm{cov}(G)$.*

The first equality was observed in [10], and the second in [9]. Both are easy to verify. If a graph $G = (V, E)$ 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, the r sets $I_i = \{u \in V : i \in A_u\}$ are independent and cover all non-edges of G, and the CNF of the form above with $S_i = V \setminus I_i$ represents the graph G.

Alon [1] used probabilistic arguments to prove that $\mathrm{cov}(G) = O(d^2 \log n)$ for every n-vertex graph G of maximum degree d. Hence, we have the following general upper bound.

Lemma 3.10 (Alon [1])**.** *For every n-vertex graph G of maximum degree d, we have $\mathrm{cnf}(G) = O(d^2 \log n)$.*

Another possibility to show that a graph H can be represented by a small monotone circuit is to design a small non-monotone circuit representing H, and then use the fact that negation is (almost) powerless in the context of graph representation.

Lemma 3.11. *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 reminiscent of the proof, due to Berkowitz [3], that negation is (almost) powerless for slice functions (see also Theorem 13.1 in [29]).

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 an equivalent 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. [29], 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. □

3.5 Lower bounds for single level circuits

Given a covering $E = \bigcup_{i=1}^{m} A_i \times B_i$ of the edges of a graph $G = (V, E)$ by bipartite cliques, its *size* is the number m of cliques, and its *weight* is the total number $\sum_{i=1}^{m}(|A_i| + |B_i|)$ of vertices in these cliques. 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 [9], and now are the subject of an extensive literature. In particular, it is known that the maximum of $\mathrm{cc}(G)$ over all n-vertex graphs is $n - \Theta(\log n)$ [6, 26, 24], and that the maximum of $\mathrm{cc_w}(G)$ is $\Theta(n^2/\log n)$ [4, 7, 5].

For a graph G, let $\mu(G)$ be the minimum of $(a + b)/ab$ over all pairs $a, b \geq 1$ such that G contains a copy of a complete bipartite $a \times b$ graph $K_{a,b}$.

Lemma 3.12. *For every graph G, $C^1_{\&}(f_G) = \mathrm{cc}(G)$ and $L^1(f_G) \geq \mu(G) \cdot |E|$. Moreover, if G is an extension of a bipartite graph H, then $\mathrm{cc}(G) \geq \mathrm{cc}(H)/2$ and $\mathrm{cc_w}(G) \geq \mathrm{cc_w}(H)$.*

Proof. The equalities $C^1_{\&}(f_G) = \mathrm{cc}(G)$ and $L^1(f_G) = \mathrm{cc_w}(G)$ follow immediately from the fact (shown in [4, 14]) that monotone single level circuits for quadratic functions have the form (3) where m is the number of AND gates in the circuit.

To show that $\mathrm{cc_w}(G) \geq \mu(G) \cdot |E|$, let $E = A_1 \times B_1 \cup \cdots \cup A_m \times B_m$ be a bipartite clique covering of $G = (V, E)$ of minimal weight. 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

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

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 $A_i \times B_i$, $i = 1, \ldots, m$ is a bipartite clique covering of G, then $(A_i \cap U) \times (B_i \cap W)$, $(B_i \cap U) \times (A_i \cap W)$, $i = 1, \ldots, m$ is a bipartite clique covering of H. The number of bipartite cliques in this new covering is at most twice that in the original covering, and the total number of vertices in the new covering does not increase at all. □

The case of circuits when we count all gates (not just AND gates) is a bit more complicated because boolean sums (entering AND gates) may not necessarily be computed separately: one partial sum computed at some OR gate may be used many times. Still, by Lemma 3.7, we know that the overlap of gates cannot be too large if the sums are disjoint enough. The disjointness of a collection of sums $\bigvee_{i \in S_1} x_i, \ldots, \bigvee_{i \in S_m} x_i$ is naturally related to the absence of large cliques in the incidence $m \times n$ graph of this collection where i and j are adjacent iff $j \in S_i$: the collection of sums is (h, k)-disjoint precisely when this graph has no copies of $K_{h+1,k+1}$. Amano and Maruoka [2] used this relation to show that $C^1(f_G) \geq |E|$ for any graph $G = (V, E)$ with no copies of $K_{2,2}$; in this case the corresponding sums are $(1, 1)$-disjoint. Their argument can be easily extended to yield a lower bound of the form $C^1(f_G) \geq |E|/t^{O(1)}$ for $K_{t,t}$-free graphs. However, we need super-linear lower bounds on $C^1(f_G)$ for graphs G which are *saturated extensions* of bipartite $n \times n$ graphs H, and such graphs already have copies of $K_{t,t}$ with $t = n/4$, even if the graph H itself is $K_{2,2}$-free.

To get rid of this problem, we use a tighter analysis of single level circuits to prove a stronger result, namely, a lower bound on the minimum size $C^1(H)$ of monotone single level circuits *representing* H (recall that such a circuit must behave correctly only on edges and non-edges of H; on other inputs it may take arbitrary values). If G is an extension of H then non-edges of H are also non-edges of G, and hence, must be rejected by f_G. This means that every circuit *computing* f_G must also represent H, implying that $C^1(f_G) \geq C^1(H)$ for every extension G of H.

Lemma 3.13. *Let $H \subseteq U \times W$ be a bipartite star-free $n \times n$ graph with no copies of $K_{t,t}$. Then $C^1(H) = \Omega(|H|/t^3)$.*

Proof. Take a minimal monotone single level circuit F representing H. The circuit F has the form $\bigvee_{i=1}^m g_i \wedge h_i$ where

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

with $S_i, T_i \subseteq U \cup W$ are boolean sums computed at the inputs of the i-th AND gate. Our goal is to show that we need many OR gates to compute these sums. We cannot apply Lemma 3.7 directly to these sums because the corresponding families may not be disjoint enough. Still, we can use the absence of $K_{t,t}$ in H to show that the restriction of these families to the left part U or to the right part W of the bipartition must contain a large enough (t, t)-disjoint subfamily.

First, observe that $S_i \cap T_i = \emptyset$ because the graph H is star-free (single variables represent complete stars). 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 (recall that on pairs of vertices within one part of the bipartition the circuit can take arbitrary values). So, we may assume that this does not happen. Hence, H is the union of bipartite cliques

$$\begin{aligned} A_i \times B_i &= (S_i \cap U) \times (T_i \cap W) \\ A_i' \times B_i' &= (T_i \cap U) \times (S_i \cap W) \end{aligned}$$

for $i = 1, \ldots, m$. We may assume w.l.o.g. that the union H' of cliques $A_i \times B_i$, $i = 1, \ldots, m$ contains at least $|H'| \geq |H|/2$ edges of H (if not, then take the remaining bipartite cliques).

Since H' 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

$$H_1 = \bigcup_{i \in I} A_i \times B_i$$

contains at least $|H_1| \geq |H'|/2 \geq |H|/4$ edges of H (if not, then let H_1 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).

This way we obtain a bipartite $K_{t,t}$-free graph $H_1 \subseteq A \times B$ with parts $A = \bigcup_{i \in I} A_i$ and $B = \bigcup_{i \in I} B_i$, and with $|H_1| \geq |H|/4$ edges. We are going to represent this graph by a monotone (single level) circuit F_1 of size not much larger than that of F, and to apply Lemma 3.7 in order to show that the size of F_1 must be large; this will yield the desired lower bound on size(F).

To achieve the first goal, we collect the boolean sums $h_i, i \in I$ computed in F into a circuit F_1, by the following construction

$$F_1(X) = \bigvee_{u \in A} x_u \wedge \Big(\bigvee_{i \in I_u} h_i \Big) = \bigvee_{u \in A} x_u \wedge \Big(\bigvee_{i \in I_u} \bigvee_{v \in T_i} x_v \Big)$$

where $I_u = \{i \in I : u \in A_i\}$. For every vertex $u \in A$, the circuit F_1 accepts an arc $uv \in A \times B$ iff $v \in T_i \cap W = B_i$ for some $i \in I$ such that $u \in A_i$. Hence, F_1 represents the graph H_1. Since all boolean sums h_i with $i \in I$ are already computed in F, 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, $\mathrm{size}(F_1) \leq \mathrm{size}(F) + 2t \cdot |I| \leq 3t \cdot \mathrm{size}(F)$.

On the other hand, by the construction, the circuit F_1 simultaneously computes all boolean sums $\bigvee_{i \in I_u} h_i = \bigvee_{v \in T_u} x_v$ with $u \in A$ and $T_u = \bigcup_{i \in I_u} T_i$ using only fanin-2 OR gates. Hence, size(F_1) is at least the disjunctive complexity of the family $\mathcal{T} = \{T_u : u \in A\}$. This, in its turn, is at least the disjunctive complexity of the restriction $\mathcal{T}' = \{T_u \cap W : u \in A\}$ of $\mathcal{T}$ to the set W: having a circuit for $\mathcal{T}$ we can get a circuit for $\mathcal{T}'$ just by setting to 0 all variables x_u with $u \notin W$. Observe that for every $u \in A$,

$$T_u \cap W = \bigcup_{i : u \in A_i} T_i \cap W = \bigcup_{i : u \in A_i} B_i$$

is the set of all neighbors of u in H_1. Since H_1 has no copies of $K_{t,t}$, no t vertices in A can have t common neighbors. This means that the family $\mathcal{T}'$ must be (t,t)-disjoint (in fact, even $(t-1,t-1)$-disjoint). Since $|H_1| = \sum_{u \in A} |T_u \cap W|$, Lemma 3.7 yields

$$\mathrm{size}(F_1) \geq \frac{1}{t^2} \sum_{u \in A} |T_u \cap W| - \frac{|A|}{t} = \frac{|H_1|}{t^2} - \frac{|A|}{t}.$$

Together with the previous estimate $\operatorname{size}(F_1) \leq 3t \cdot \operatorname{size}(F)$ and an obvious estimate $|A| \leq t \cdot |I| \leq t \cdot \operatorname{size}(F)$, this yields

$$\operatorname{size}(F) \geq \frac{1}{3t} \cdot \operatorname{size}(F_1) \geq \frac{|H_1|}{3t^3} - \frac{|A|}{3t^2} \geq \frac{|H_1|}{3t^3} - \operatorname{size}(F).$$

Since $|H_1| \geq |H|/4$, the desired lower bound $\operatorname{size}(F) = \Omega(|H|/t^3)$ follows. □

Now we turn to the actual proof of Theorems 2.1–2.4.

4 Circuits: proof of Theorem 2.1

In order to prove the gap, claimed in Theorem 2.1, we need (by Lemma 3.13) a bipartite $n \times n$ graph which

1. is dense, i.e., has $\Omega(n^2)$ edges,
2. has no copies of $K_{t,t}$ with t about $\log n$,
3. can be represented by a small (linear size) monotone circuit.

The existence of graphs, satisfying the first two conditions, is a classical result of Erdős [8]. However, its proof is probabilistic and gives no idea on how to ensure the third condition. To get rid of this problem, we just reverse the order of the argument: we first choose an appropriate graph G whose induced subgraphs satisfy the third condition. Then we use the probabilistic argument to show that G must contain a sufficiently large induced subgraph satisfying the first two conditions.

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 4.1 (Pudlák–Rödl [20]). *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 H_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 $y \in Y$. Viewing the vectors in X as the rows of the coefficient matrix and the 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.

In the next lemma we use the following versions of Chernoff's inequality (see, e.g., [18], § 4.1): if X is the sum of n independent Bernoulli random variables with the success probability p, then $\Pr(|X| \leq (1-c)pn) \leq e^{-c^2pn/2}$ for $0 < c \leq 1$, and $\Pr(|X| \geq cpn) \leq 2^{-cpn}$ for $c > 2e$.

Lemma 4.2. *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 $K_{r,r}$.*

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 $p|V| = 2$ elements in $|\mathbf{S} \cap V|$. By Chernoff's inequality, $\Pr\left(|\mathbf{S} \cap V| \geq 2c\right) \leq 2^{-2c}$ holds for any $c > 2e$. The number of vector spaces in $\mathbb{F}^r$ of dimension $r/2$ does not exceed $\binom{r}{r/2} \leq 2^r/\sqrt{r}$. We can therefore take $c = r/2$ and conclude that the set $\mathbf{S}$ intersects some $r/2$-dimensional vector space V in $2c = r$ or more elements with probability at most $2^{r-(\log r)/2-r} = r^{-1/2} = o(1)$. Hence, with probability $1 - o(1)$ the set $\mathbf{S}$ has cardinality at least $2^{r/2}$ and $|\mathbf{S} \cap V| < r$ for every $r/2$-dimensional vector space V. Fix such a set S' and take an arbitrary subset $S \subseteq S'$ of cardinality $|S| = 2^{r/2}$. By Lemma 4.1, neither H_S nor $\overline{H}_S$ contains a copy of $K_{r,r}$. □

Now we turn to the actual proof of Theorem 2.1.

Proof of Theorem 2.1. Let $S \subseteq \mathbb{F}^r$ be a subset of cardinality $|S| = n = 2^{r/2}$ guaranteed by Lemma 4.2. 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 $|H| \geq |S|^2/2 = n^2/2$ edges. Moreover, this graph contains no copy of $K_{r,r}$ where $r = 2\log n$.

Let now G be the saturated extension of H. By removing the centers of complete stars, we obtain an *induced* star-free subgraph H' of H. Since the graph H has no copies of $K_{r,r}$, it can have at most $2(r-1)$ complete stars, implying that the resulting subgraph H' still has $|H'| \geq |H| - 2(r-1)n = \Omega(n^2)$ edges. Moreover, every circuit representing H must also represent H', just because edges/non-edges of H' are also edges/non-edges of H (this is a property of *induced* subgraphs, not shared by spanning subgraphs) and every circuit for H must correctly accept/reject them. Therefore, $C^1(f_G) \geq C^1(H) \geq C^1(H')$ where, by Lemma 3.13, $C^1(H') = \Omega(|H'|/r^3) = \Omega(n^2/\log^3 n)$. Hence, $C^1(f_G) = \Omega(n^2/\log^3 n)$.

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 verify 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 3.6, all $r = 2\log n$ boolean sums in the formula $F(X)$ above can be simultaneously computed by a circuit of linear (in n) size. Hence, the graph H can be represented by a linear size circuit over the basis $\{\vee, \wedge, \neg\}$ and, by Lemma 3.11, can be represented by a *monotone* circuit of linear size. Since G is the saturated extension of H, Lemma 3.8 implies that $C(f_G) = O(n)$. Hence, $\mathrm{Gap}(G) = C^1(f_G)/C(f_G) = \Omega\left(n/\log^3 n\right)$.

5 Formulas: proof of Theorem 2.2

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 $|K| = \sum_{u \in U} 2^{r-|u|} = 3^r \geq n^{3/2+c}$ 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 that $(\bigcup_{u \in A} u) \cap (\bigcup_{v \in B} v) = \emptyset$. Since $a \leq a'$ and $b \leq b'$ imply $(a+b)/ab \geq (a'+b')/a'b'$, we have $\mu(K) \geq (2^k + 2^{r-k})/2^r \geq 2^{-r/2} = n^{-1/2}$. By Lemma 3.12, $L^1(f_G) = \mathrm{cc_w}(G) \geq \mathrm{cc_w}(K) \geq \mu(K) \cdot |K| \geq n^{1+c}$.

On the other hand, by its definition, the graph K admits an intersection representation of size r and, by Lemma 3.9, can be represented by a monotone CNF with $\mathrm{int}(K) \leq r = \log n$ clauses, and hence, by a monotone formula with $O(n \log n)$ fanin-2 AND and OR gates. Together with Lemma 3.8, this implies that $L(f_G) = O(n \log n)$. Hence, $\mathrm{Gap}_{\mathrm{form}}(G) = L^1(f_G)/L(f_G) = \Omega(n^c/\log n)$.

6 Unbounded fanin circuits: proof of Theorem 2.3

To prove the lower bound $C_*(f_G) \geq \mathrm{m}(G) + 1$ we use the communication complexity argument. By an observation due to Nisan (see [19] or [13], Lemma 11.2), $C_*(f_G)$ is at least the deterministic two-party communication complexity of f_G under the worst-case partition of its input variables (this holds for arbitrary, not necessarily quadratic, functions and for arbitrary, not necessarily monotone, circuits). Let now M be an induced matching in G with $|M| = \mathrm{m}(G)$ edges. By setting to 0 all the variables corresponding to vertices outside this matching, we obtain that $C_*(f_G) \geq C_*(f_M)$ (recall that M is an *induced* subgraph of G). The function f_M itself has the form $f_M = \bigvee_{i=1}^{|M|} x_i y_i$, i.e., is the negation of the set disjointness function, and its deterministic communication complexity under the natural partition where one player gets all x_i's and the other gets all y_i's is well known to be $|M| + 1$. Hence, $C_*(f_G) \geq C_*(f_M) \geq |M| + 1 = \mathrm{m}(G) + 1$.

For the proof of the second part of Theorem 2.3 we need the following fact. Let $\mathrm{cnf}(f_G)$ denote the minimum length of (i.e. the number of clauses in) a monotone CNF computing f_G.

Lemma 6.1. *For every graph G of maximum degree d, $\mathrm{cnf}(f_G) \geq \tau(G)/d$.*

Proof. Let F be a monotone CNF of length $t = \mathrm{cnf}(f_G)$ computing f_G. Since f_G has no prime implicants of length 1 (by its definition (1)), 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 G, each of these edges must have at least one endpoint in S. Hence, S must be a vertex cover of G, implying that $|S| \geq \tau(G)$.

Since F is a shortest CNF computing f_G, the shrinked CNF F' must make an error, i.e. it must (wrongly) accept some independent set of G. 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 G) with at least one vertex in I, for otherwise F would wrongly accept the independent set $I \cup \{u\}$ of G. Hence, at least one vertex $v \in I$ must have at least $|S|/|I| \geq \tau(G)/t$ neighbors in S. Since the degree of v cannot exceed d, the desired lower bound $t \geq \tau(G)/d$ follows. □

Take now an arbitrary graph $G=(V,E)$ of maximum degree d, and let F be a smallest monotone Σ_3 circuit computing f_G. We first consider the case when F is a formula, i.e. all gates have fanout 1. 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 $E_1 \cup \cdots \cup E_s = E$. Note that $\tau(G) \leq \sum_{i=1}^{s} \tau(G_i)$. Since each of these subgraphs has maximum degree at most d, Lemma 6.1 implies that the entire formula F must have size at least $\sum_{i=1}^{s} r_i \geq \sum_{i=1}^{s} \tau(G_i)/d \geq \tau(G)/d$. If F is not a formula (some OR gates on the bottom level have fanout larger than 1), then we still have that $\mathrm{size}(F) \geq t = \max\{s, r_1, \ldots, r_s\}$. Take a CNF F_i for which $\tau(G_i) \geq \tau(G)/s$. By Lemma 6.1, F_i has length $r_i \geq \tau(G_i)/d \geq \tau(G)/sd$. Since both r_i and s do not exceed t, this yields $t^2 \geq \tau(G)/d$, and the desired lower bound $t \geq \sqrt{\tau(G)/d}$ on the number of gates in F follows.

7 Multiplicative complexity: proof of Theorem 2.4

Let G be the saturated extension of an n to n matching M. Then, by Lemma 3.12, $C^1_{\&}(f_G) = \mathrm{cc}(G) \geq \mathrm{cc}(M)/2 = n/2$. On the other hand, M can be represented by a monotone CNF of length $O(\log n)$. This follows from a more general Lemma 3.10, but can also be shown directly (see [10]): 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 of A_i. It is clear that then $A_i \cap B_j = \emptyset$ iff $i=j$. Hence, $\mathrm{cnf}(M) = \mathrm{int}(M) \leq r = 2\log n$. Together with Lemma 3.8, this implies that $C_{\&}(f_G) = O(\log n)$. Hence, $\mathrm{Gap}_{\mathrm{mult}}(G) = \Omega(n/\log n)$.

8 Concluding remarks and open problems

As we mentioned in § 2, the unbounded fanin version of the single level conjecture is true for almost all graphs. Better yet, Theorem 2.3 implies that the conjecture is true for all n-vertex graphs containing an induced matching with $\Omega(n)$ edges. Still, it seems very unlikely that the conjecture is true for all graphs.

Problem 8.1. *Does there exist n-vertex graphs G of maximal degree d with $L_*(f_G) = o\left(\tau(G)/d\right)$ or $C_*(f_G) = o(\sqrt{\tau(G)/d})$?*

A next open question is to prove *super-linear* lower bounds on the size of monotone (fanin 2) circuits computing explicit quadratic functions in n variables. For formulas (fanout 1 circuits) lower bounds $L(f_G) = \Omega(n^{3/2})$ can be proved using the rank argument

[10]. However, the case of *circuits* is more complicated because (as mentioned in the introduction) known lower bounds for monotone circuits—the method of approximations due to Razborov [22], and its derivatives—cannot yield lower bounds larger than n.

Problem 8.2. *Prove $C(f_G) \geq n^{1+\epsilon}$ for an explicit n-vertex graph G.*

What can be said about the single level conjecture in the context of graph representation, that is, if we consider circuits *representing* graphs G instead of circuits *computing* their quadratic functions f_G? For circuits with fanin-2 gates the question is already answered in § 4: the gap between single level and general circuits is $\Omega(n/\log^3 n)$ also in this context. But what about circuits with unbounded fanin gates? For a graph G, let $C_*(G)$ be the minimum size of a monotone unbounded fanin circuit representing G, and let $C_*^1(G)$ be the single level version of this measure. Note that, for some graphs G, circuits representing G may be exponentially smaller than circuits computing the quadratic function f_G. If, say, M_n is a matching with n edges, then $\mathrm{cnf}(M_n) = O(\log n)$ (by Lemma 3.10) but $C_*(f_{M_n}) = \Omega(n)$ (by Theorem 2.3). This also shows that, in the context of graph representation, Lemma 6.1 does not hold anymore.

Problem 8.3 (Pudlák–Rödl–Savický [21]). *Prove that $C_*^1(G)$ may be much larger than $C_*(G)$.*

Easy counting shows that $C_*^1(G) = \Omega(n)$ for almost all n-vertex graphs. On the other hand, as mentioned in § 2, a lower bound $n^{\Omega(1)}$ for an explicit graph G would yield a super-linear lower bound for non-monotone log-depth circuits. Actually, even a much more moderate lower bound $2^{\alpha\sqrt{\log n}}$ with $\alpha \to \infty$ would have interesting consequences (see [10]).

Problem 8.4. *Prove $C_*^1(G) \geq 2^{\alpha\sqrt{\log n}}$ for an explicit n-vertex graph G.*

Although, as mentioned above, we already can prove lower bounds $L(f_G) = \Omega(n^{3/2})$ for some explicit graphs G, doing this for *saturated* graphs is a much more difficult task. Bloniarz [4] used counting arguments to show that $C(f_G) = \Omega(n^2/\log n)$ for almost all n-vertex graphs G; this remains true also in the class of saturated graphs. The problem, however, is the *explicitness*: we want a lower bound for explicitly constructed graphs. As mentioned in the introduction, a lower bound $C(f_G) \geq cn$ for a sufficiently large constant $c > 0$ would have great consequences in circuit complexity. A (potentially) less ambitious problem is to do this for formulas.

Problem 8.5. *Exhibit an explicit saturated star-free graph on n vertices with $L(f_G) = \Omega(n \log^k n)$.*

Since, by Observation 3.5, for such graphs we have the equality $L(G) = L(f_G)$, this would yield an explicit boolean function in $m = \Theta(\log n)$ variables requiring non-monotone formulas of size $\Omega(m^k)$ (see [10] for details).

Acknowledgments

I am grateful to Georg Schnitger and Ingo Wegener for interesting discussions.

References

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

[2] M. Amano and A. Maruoka, *On the monotone circuit complexity of quadratic Boolean functions*, in Proc. of 5-th Int. Symp. on Algorithms and Computation, Springer Lect. Notes in Comput. Sci., vol. 3341 (2004), pp. 28–40.

[3] S.J. Berkowitz, *On some relations between monotone and non-monotone circuit complexity*, Tech. Rep., Comput. Sci. Dept., University of Toronto, 1982.

[4] P. A. Bloniarz, *The complexity of monotone boolean functions and an algorithm for finding shortest paths in a graph*, PhD Dissertation, Tech. Rep 238, Lab. Comput. Sci., MIT, Cambridge, MA, 1979.

[5] S. Bublitz, *Decomposition of graphs and monotone size of homogeneous functions*, Acta Inform., 23 (1986), pp. 689–696.

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

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

[8] P. Erdős, *Some remarks on the theory of graphs*, Bull. Amer. Math. Soc., 53 (1947), pp. 292–294.

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

[10] S. Jukna, *On graph complexity*, ECCC Tech. Rep. Nr. 5, 2004. To appear in Combin. Probab. Comput.

[11] M. Kneser, *Aufgabe 300*, Jahresber. Deutsch. Math.-Verein, 58(2) (1955), p. 27.

[12] R. E. Krichevski, *Complexity of contact circuits realizing a function of logical algebra*, Sov. Phys. Dokl., 8 (1964), pp. 770–772.

[13] E. Kushilevitz and N. Nisan, *Communication Complexity*, Cambridge University Press, 1997.

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

[15] L. Lovász, *Kneser's conjecture, chromatic numbers and homotopy*, J. Comb. Th. (A), 25 (1978), pp. 319–324.

[16] K. Mehlhorn, *Some remarks on Boolean sums*, Acta Inform., 12 (1979), pp. 371–375.

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

[18] R. Motwani and P. Raghavan, *Randomized Algorithms*, Cambridge University Press, New York, 1995.

[19] N. Nisan, *The communication complexity of threshold gates*, in Combinatorics, Paul Erdős is Eighty, vol. 1, D. Miklós, V.T. Sós and T. Szőni, Eds., Janos Bolyai Math. Society, Budapest, Hungary (1993), pp. 301–315.

[20] P. Pudlák and V Rödl, *Pseudorandom sets and explicit constructions of Ramsey graphs*, in J. Krajiček (ed.) Quad. Mat., 13 (2004), pp. 327–346.

[21] P. Pudlák, V. Rödl, and P. Savický, *Graph complexity*, Acta Inform., 25 (1988), pp. 515–535.

[22] A. Razborov, *Lower bounds on the monotone complexity of some Boolean functions*, Soviet Math. Dokl., 31 (1985), pp. 354–357.

[23] A. Razborov and S. Rudich, *Natural proofs*, J. Comput. Syst. Sci., 55(1) (1997), pp. 24–35.

[24] V. Rödl and A. Ruciński, *Bipartite coverings of graphs*, Combin. Probab. Comput., 6 (1997), pp. 349–352.

[25] R. Tarjan, *Complexity of monotone networks for computing conjunctions*, Ann. Discrete Math., 2 (1978), pp. 121–133.

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

[27] L. Valiant, *Graph-theoretic methods in low-level complexity*, in Proc. of 6-th Conf. on Math. Foundations of Comput. Sci, Springer Lect. Notes in Comput. Sci., vol. 53 (1977), pp 162–176.

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

[29] I. Wegener, *The Complexity of Boolean Functions*, Wiley-Teubner, 1987.