58th Annual IEEE Symposium on Foundations of Computer Science

On small-depth Frege proofs for Tseitin for grids

Johan Håstad
School of Computer Science and Communication
KTH Royal Institute of Technology
Stockholm, Sweden
Email: johanh@kth.se

***Abstract*—We prove a lower bound on the size of a small depth Frege refutation of the Tseitin contradiction on the grid. We conclude that polynomial size such refutations must use formulas of almost logarithmic depth.**

***Keywords*-Frege proofs; proof complexity; switching lemma, Tseitin formulas;**

I. Introduction

This paper is in the setting of propositional proof complexity. We are given a propositional statement and some reasoning rules. The most basic proof system is resolution. In this proof system we study clauses, i.e. disjunctions of literals and have a simple way to derive new clauses from existing clauses. If we derive the empty clause we have reached a contradiction refuting the original formula.

Resolution has been studied extensively and by now we have a large body of work understanding the strengths and limitations of resolution. In an early paper [1], Tseitin defined the set of contradictions based on graphs studied in this paper and proved that any regular resolution proof of this contradiction requires exponential size proofs in general. A later result by Haken [2] gave the first strong lower bound for unrestricted resolution proving that the pigeon-hole principle (PHP) requires exponential size proofs. As this paper is not about resolution let us not discuss the many strong results obtained, but only mention the famous paper of Ben-Sasson and Wigderson [3] as a high point which among other results showed the importance of width when studying resolution proofs.

There are many proof systems that are more powerful than resolution and in this paper we study the case when each formula appearing in the proof is restricted to be a Boolean formula of small depth d. Here $d = 1$ corresponds to resolution. There are many alternatives for the reasoning rules and what is said below applies to any constant size set of reasoning rules that are consistent. The first strong result was obtain by Ajtai [4] showing that the PHP cannot be proved in constant depth and polynomial size.

Ajtai did not work out an explicit lower bound for the depth of polynomial size proofs but in a later reformulation by Bellantoni et al. [5], a lower bound of $\Omega(\log^* n)$ was given. This was later strengthened [6], [7] to obtain $\Omega(\log\log n)$ lower bounds for PHP. Similar bounds were later proved by Urquhart and Fu [8] and Ben-Sasson [9] for Tseitin formulas for the complete graph and for constant-degree expander graphs, respectively.

On the positive side Buss [10] proved that there are polynomial size $O(\log n)$-depth proofs for the PHP and his methods can be adopted to also yield similar proofs for the Tseitin formulas for any constant-degree graph.

The exponential gap between the depth bounds $\log\log n$ and $\log n$ was recently partly closed by Pitassi et al. [11] obtaining a $\Omega(\sqrt{\log n})$ lower bound for Tseitin formulas on a certain 3-regular expander graph. It is curious to note the the size lower bounds of [11] when considering depth d is exponential in $\Omega((\log n)^2/d^2)$ and thus only weakly superpolynomial. For small values of d, this bound is weaker then the bounds of the form $exp(n^{c^{-d}})$ obtained in previous paper but extends the range of d for which it is superpolynomial.

In the current work we study the Tseitin formulas for the 2-dimensional grid and almost close the gap obtaining size lower bounds $exp(\Omega(n^{1/58(d+1)}))$ for depth d proofs and hence the depth lower bound $\Omega(\log n/\log\log n)$ for polynomial size proofs. Our proofs follow the same paradigm as earlier proofs and let us sketch the underlying mechanisms at a semi high level to put our contribution in perspective.

When studying circuits of small depth it has turned out to be profitable to study restrictions that fix most of the input variables to constants. This is useful as for suitably chosen restrictions it is possible to decrease the depth of almost all small circuits by one. This was first used to prove lower bounds for circuits size [12], [13], [14], [15] and the simplest case is when proving lower bounds for the size of depth-d circuits computing parity and let us briefly discuss this case.

In this situation one uses the simplest space of random restrictions usually denoted by R_p. In such a restriction each input variable is kept with probability p and otherwise set to 0 or 1 with equal probabilities. The key notion for decreasing depth is a switching lemma which says that if you are given a depth two circuit with bottom fanin t then, if you at the same time apply a restriction, it can be switched to a depth 2 circuit of the other type of bottom fanin s, except with probability at most $(5pt)^s$.

Using this switching property for the two layers closest to the inputs creates two adjacent layers of gates of the same

0272-5428/17 $31.00 © 2017 IEEE
DOI 10.1109/FOCS.2017.18

type which makes it possible to decrease the depth of the circuit by one. To prove a lower bound for parity one just needs to make the trivial observation that the resulting circuit must compute the parity (or its negation) of the remaining variables. Applying $d-1$ restrictions we are able to make the circuit simple enough to be analyzed directly. The number of remaining variables is about $p^{d-1}n$ and we simply need a large enough p to make this number non-trivial.

To prove lower bounds for the size of proofs for various families of formulas one needs more subtle restrictions. We are no longer computing a function but instead given a set of axioms. We want that a restriction reduces the problem to a smaller problem of the same type. This is more or less equivalent to that each axiom is either reduced to an axiom of the smaller instance or to something that is a tautology. We must, at all cost, make sure that no axiom is made false as this would imply that the contradiction we are trying to prove cannot be produced efficiently is available as an axiom. In most cases each axiom is of constant size and this implies that we cannot use spaces of restrictions, such as R_p, that treat the variables independently. Restrictions that give values in a dependent way cause problems with the proof (or even validity) of the switching lemma. The key is thus finding a balance between the property of preserving the axioms of the formula we are studying while still being able to prove a switching lemma with good parameters.

On the high level, the strength of a switching lemma is controlled by the size of the obtained problem (which corresponds to the parameter p for independent restrictions) and how the failure probability depends on the parameters s and t. The switching lemma of [11] has failure bounds on the form $(cpt2^t)^s$. The reason for this factor 2^t is a bit mysterious and [11] conjectures that it is not needed. The paper by Mehta [16] describes similar situations where the factor is indeed needed. The worse bounds force the proof to work with small values of s and t (roughly $(\log n)/d$) and a p that is about 2^{-ct} for a $c > 1$.

We improve the troublesome factor 2^t of [11] to the better, but probably not optimal, factor t^c for a constant c. This implies that the loss in one application of the switching lemma roughly corresponds to c applications of the lemma with the optimal parameters and thus we get this multiplicative factor in front of d. As this is a constant we get asymptotically sharp bounds for the depth of polynomial size proofs.

A key point in the proof is the choice of the space of restrictions. The high level picture is not that surprising. Given a $n \times n$ grid we pick sub-squares of size $T \times T$ (where T is poly-logarithmic when studying polynomial size proofs and $n^{\Omega(1/d)}$ in general) and in each sub-square we pick a node and connect the picked nodes by paths. For each path P we have a new variable x_P and for each edge e on the path it is either replaced by x_P or its negation $\bar{x}_P$. This is done in a way that irregardless of the values of these new variables all constraints, except at the picked nodes are automatically satisfied while the constraints at the picked nodes give the constraints of the smaller instance. The essential new part of the current paper is the choice of restrictions and the proof of the switching lemma. The way to analyze how restrictions make all sub-formulas be represented by small-depth decision trees is similar to previous papers.

An overview of the paper is as follows. We start with some preliminaries in Section II and proceed with some properties of the grid and assignments that satisfies some parity conditions in Section III. We define our restriction in Section IV. We proceed to recall the formalism of t-evaluations in Section VI after having described some basic properties of consistent decision trees in Section V. Assuming the switching lemma we are able to complete the proof of our main theorem also in Section VI and we end by the proof of the switching lemma in Section VII.

II. Some preliminaries

The Tseitin contradiction for a graph G is a statement on a set of variables x_e where e ranges over all edges of G. It is usually stated for a small degree graph and in our case G is what is usually, and also here, called "the grid" but in fact it is the torus. In other words we have nodes indexed by (i, j), for $0 \leq i, j \leq n-1$ where n is an odd integer and a node (i, j) is connected to the four nodes at distance 1, i.e. where one coordinate is identical and the other has moved up or down by 1 modulo n. The constraint at node v is that the number of edges, e, incident to v with $x_e = 1$ is odd. This is a contradiction as can be seen by summing the constraints over all nodes. Each edge appears in two equations while the right hand sides sum to 1 modulo 2. The contradiction can be formulated as a 4-CNF formula by having 8 clauses of length four for each node.

Our formulas only contain $\vee$-gates and negations. We simulate $\wedge$ in the obvious way and depth is the number of alternations of $\vee$ and $\neg$.

III. Properties of assignments on the grid and dynamic matchings

We are interested in assignments to some of the variables on the grid and to what extent they satisfy the parity constraints that the variables around any vertex sum to 1 modulo 2. Of course we cannot have a total assignment that satisfies all constraints simultaneously but we have plenty of assignments that satisfy all constraints in some particular area of the grid.

On a set X of nodes we say that a partial assignment is *complete* if it gives values to exactly all variable with at least one endpoint in X. The support of a partial assignment is the set of nodes adjacent to a variable given a value. Note that the support of a complete assignment on X also includes the neighbors of X.

In our proof we consider assignments to small sets of variables and in particular we are interested in cases where the size of the set X is $o(n)$ and hence a small part of the grid. Let X^c denote the complement of X.

In the situation when X is small, X^c contains a giant component containing almost all nodes of the grid. This follows as there are at least $n - o(n)$ complete rows and columns in X^c and the nodes of these rows and columns are all connected. For a set X let the closure of X, $c(X)$ denote all nodes either in X or in small connected components of X^c. Note that $c(X)^c$ is exactly the giant component of X^c.

Definition 3.1: An assignment α supported on a set X is *locally consistent* if it can be extended to a complete assignment on $c(X)$ that satisfies all parity constraints on this set.

We extend this definition to say that two assignments are consistent with each other if they do not give different values to the same variable and when you look at the union of the two assignment this gives a locally consistent assignment. Let us prove a simple but useful lemma.

Lemma 3.2: Suppose α is a locally consistent assignment supported on a set of size $o(n)$ and x_e a variable not in the support of α. Then there is a locally consistent assignment α' that extends α and gives a value to x_e.

Proof: Suppose the support of α is X and let X^+ be X with the endpoints of e added. First extend α to be an assignment that satisfies the constraints on $c(X)$ and then take any further extension that gives values to all variables touching $c(X^+)$. Suppose this assignment violates the parity constraint at a node v. Take a path that starts at v and ends in $c(X^+)^c$ and does not pass through any node in $c(X)$. This is possible as $c(X)^c$ is connected. Negate the variables corresponding to edges on this path. The new assignment satisfies the constraint at v, still extends α and does not cause any new violations on $c(X^+)$. Repeating this procedure for any $v \in c(X^+)$ that has its constraint violated creates a locally consistent assignment that extends α and gives a value to x_e. ∎

The technique used in this proof of taking a path between two nodes and flipping the values along this path is extremely useful for thinking about assignments under the Tseitin conditions. This changes the validity of the constraints at the endpoints but preserves the constraints at all other nodes. Next we discuss a dynamic matching game needed by our analysis.

We have two players, one adversarial player that supplies nodes while the other, matching player P_M, is supposed to dynamically create a matching that contains the nodes given by the adversarial player. As the full grid is of odd size and hence does not have a perfect matching the adversarial player will eventually win, but clearly P_M can survive for a while and this is sufficient for us. To be more precise we have the below lemma proved in the full version of the paper.

Lemma 3.3: When the dynamic matching game is played on the $n \times n$ grid, P_M can survive for at least $n/2$ moves.

IV. Restrictions

The plan is to make a probabilistic assignment to variables of the grid that reduces the Tseitin contradiction to a smaller contradiction of the same type in a way that enables us to simplify all formulas appearing in an attempted proof. As the final product is a rather rigid object we utilize an intermediate partial restriction that leaves slightly more variables unset but has better independence properties. We start by defining the full restrictions.

A. Full restrictions

In an $n \times n$ grid we make sub-squares of size $T \times T$ where we assume that the number of sub-squares is odd. In each sub-square we choose[1] $\Delta = \sqrt{T}/2$ of the nodes and call them *centers*. These are located evenly spaced on the diagonal of the $3T/4 \times 3T/4$ central sub-square. This implies that they have separation $3\sqrt{T}/2$ in both dimension.

The centers in neighboring sub-squares are connected by paths that are edge-disjoint except close to the endpoints. Let us describe how to connect a given center to a center in the sub-square on top, the sideways case being analogous. There are $T/4$ rows between the two central areas. For each pair of centers (one in the top sub-square and one in the bottom) we can hence designate a unique row in this middle area.

Now for the jth center in the lower sub-square to connect to the ith center above we first go i steps to the right and then straight up to the designated row. This is completed by starting at the upper center and then going j steps to the left and then down to the designated row. We finally use the appropriate segment from the designated row to complete the path.

These paths have the property that the Δ first and last edges belong to several paths, always starting at the same center and going in the same direction while the rest of the edges on a path uniquely identifies the entire path and hence both endpoints. The "direction" is here and elsewhere in the paper counted as the relative position of the sub-squares and is thus one of "up", "down", "left" or "right". It is important for us that for any edge there is a unique center that is the closest endpoint of all paths going through this edge and all paths that go through that edge are in the same direction.

A restriction is defined by first picking one center in each $T \times T$ sub-square and then the paths described above connecting these centers. Note that these paths are edge-disjoint. The picked centers naturally form a $n/T \times n/T$ grid if we interpret the paths between the chosen centers as edges.

[1]For simplicity we assume that some arithmetical expressions that are supposed to be integers are in fact exact as integers. By a careful choice of parameters this can be achieved but for the time being we leave this detail to the reader.

We proceed to make the correspondence more complete by assigning values to variables.

Each variable is given a value such that, at any node which is not chosen, the parity of the sum of the adjacent variables is one while the same parity at chosen centers is zero. As the number of chosen centers is odd there is such an assignment. For variables not on the chosen paths these are the final values while for variables on the chosen paths we call them *suggested* values.

For each path P between two chosen centers we have a new variable x_P and for each variable x_e on the path it is replaced by x_P if the suggested value of x_e is 0 and otherwise it is replaced by $\bar{x}_P$.

We claim that with these substitutions we have reduced the Tseitin problem on an $n \times n$ grid to the same problem on an $n/T \times n/T$ grid. This is true in the sense that we have an induced grid when we interpret paths as new edges and we need to see what happens to the axioms.

Given a formula F we can apply a restriction to it in the natural way. Variables given constant values are replaced by constants while surviving variables are replaced by the appropriate negation of the corresponding path variable. A restriction has a natural effect on the Tseitin contradiction as follows.

- The axioms for nodes not on a chosen paths are all reduced to true as all variables occurring in them are fixed in such a way that the axioms are true.
- The axioms for interior nodes of a chosen path are reduced to tautologies as the axiom is true independent of the value of the involved variable(s) x_P. This is true as flipping a single x_P changes the value of two variables next to any such node.
- The axioms at the chosen centers turn into the axioms of the smaller instance.

These just defined restrictions are called *full restrictions* as they completely reduce a full size problem to a smaller problem. A typical full restriction is denoted by σ. We construct such a full restriction by first making a partial restriction and we turn to defining these next.

B. Partial restrictions and pairings

A typical partial restriction is called ρ and as we mostly discuss partial restrictions we simply call the them restrictions while we use the term "full restrictions" when that is what we have in mind. At the same time as describing partial restrictions we give a probability distribution on such restrictions.

Let k be an odd integer of the form $Cs(n/T)^2$ for a constant C. The first step of constructing ρ is picking k centers uniformly at random from the set of all $\Delta(n/T)^2$ centers. These are the *alive* centers. In the future we only consider the case when the number of live centers in each sub-square is between a factor .99 and 1.01 of its expected value Cs. The probability of this being false is $O(n^2 e^{-\Omega(s)})$ and this is simply added to other failure probabilities.

For all variables not on any path between live centers we fix these to random values respecting the parity constraints at these nodes.

We now randomly pick preferred values for all remaining variables. These preferred values satisfy all parity constraints, *except* at the live centers where they all violate the parity constraint. As the number of live centers is odd, there is one, and indeed many, such assignments of fixed and preferred values.

The choice of the centers together with the fixed and preferred variables is denoted by ρ. The choice of ρ is the main probabilistic event. Note that the number of possible values for fixed and preferred values is independent of which centers are alive and even of k as long as it is odd. This is true as the values (fixed and preferred) are selected to satisfy a number of linear equations. The left hand sides of these equations are always the same as we sum variables over all edges adjacent to a node while the right hand sides do change. Any choice for the right hand side for which there is some solution has the same number of solutions as this only depends on the number of variables and the rank of the linear system of equations.

A partial restriction ρ is, for the analysis, preferable to a full restriction σ as it behaves more independently. A drawback is, however, that as soon as a live center v is discovered then we have many paths leaving v in ρ and this could result in a deep decision tree if they all corresponded to a live variable. In order to avoid this we add a second step, a pairing π, turning a partial restriction into a full restriction.

Choose one center to survive in each sub-square. These are called the *chosen* centers and paths between such centers correspond to the variables that remain and are called chosen paths. Centers that were alive through the first part of the process but are not chosen are called *non-chosen*. The centers killed already by ρ are simply called dead.

The simplest way to eliminate the non-chosen centers would be if we were able to pair them up in such a way that the two centers in a pair are in adjacent sub-squares and hence connected by a path. In such a case we could negate the preferred values along any such path and after this make the preferred values permanent outside the chosen paths. For variables on the chosen paths we turn the preferred values into suggested values completing the full restriction.

It might be that there is such a pairing with high probability but we allow a more general way of eliminating non-chosen centers as this is easier to find. We still call the process a pairing as it is not too far from the truth and gives the right intuition.

Let us consider a graph on the non-chosen centers where two centers are connected if there is a path between them (which is true iff they are in neighboring sub-squares).

As the original grid is also a graph with edges we from

now on use the term "grid-edges" to refer to edges in the original grid. The chosen centers form a smaller grid and this also has edges and we call these "new grid-edges". We only consider paths in the grid and we keep the shorter term "path" for these. Thus from now on an "edge" is a connection between two live centers and corresponds to a path in the grid-graph. A "new grid-edge" corresponds to a chosen path and is thus also an edge in the graph of the live centers.

Some edges are conflicting in that they cannot be present in the graph at the same time. This is because we allow at most one path in each of the four directions from a center.

Our second part of the full restriction is an odd degree sub-graph π that covers all non-chosen centers such that each connected component of π is either an edge or a star with four nodes. A proof of the below lemma is given in the full paper.

Lemma 4.1: If each sub-square has between $.99Cs$ and $1.01Cs$ non-chosen centers such a pairing π exists.

As stated above π makes it possible to turn ρ into σ. Variables not on live paths take their fixed values. Variables on live paths but not on chosen paths take their preferred values unless they are on a path chosen by π in which case these values are negated. On the chosen paths the preferred values now becomes suggested and this completes the description of σ.

We use the term "preferred values" as a vast majority of the variables will eventually be fixed to these values as very few variables appear on the paths of π. On the other hand "suggested values" are much less certain as the path variables should be thought of as equally likely to be 0 and 1 and thus these variables are equally likely to take also the non-suggested value.

As an intermediate between ρ and the full restriction σ we have ρ and some information in the form of edges and "non-edges" which says that there is no edge from a certain center in a certain direction. We call such an edge or non-edge a *piece of information* and we let an *information set* I be a collection of pieces of information. An information set is consistent if it does not have two different pieces of information from the same center in one fixed direction and furthermore, if it has the information in all four directions from a center v then it has an odd number of edges. Note that here, as opposed to the grid, we do not have a problem of small connected components in the complement of a set of nodes. A center has a potential edge to all centers in neighboring sub-squares and thus this is much more connected graph than the grid. We need the notion of a closed information set.

Definition 4.2: An information set I is *closed* if it is supported on a set X of centers such that for any $v \in X$ the set I contains the information in all four directions and any edge in I is between two centers of X.

It follows that the size of X must be even and I contains a non-edge from any v in a direction where X does not have an element. When considered as a graph such an information set is an odd-degree graph (with degrees one and three) on the centers of X. One more definition.

Definition 4.3: Let ρ be a restriction and I an information set. A variable x_e is considered *forced by* (ρ, I) iff either its closest endpoint, v, is not live in ρ or if the information of v in the direction of e is contained in I. It is forced to its preferred value unless the information states that there is an edge from v in the direction of e that corresponds to path that passes through e in which case it takes the opposite value.

There are other situations where the value of a variable might be determined by ρ and I, such as the lack, or scarcity, of live centers in a sub-square but it is not allowed to use this information.

Note that if we have a closed information set I then if we consider all variables forced by (ρ, I) this can be described by a restriction where the centers in the support of I are killed. We simply negate the values of any preferred variable on any path in I and then forget that the centers in the support of I were alive.

If we let such a closed information set operate on a restriction ρ we get a restriction with fewer live centers where the number of killed centers is exactly the number nodes in the support of the corresponding graph.

V. Decision trees

We have decision trees where each internal node is marked with a variable x_e and the outgoing edges are marked with 0 and 1. The leaves of a decision tree are labeled by 0 and 1.

All decision trees considered in this paper have a depth that is smaller than the dimension of the grid we are currently considering. For each path in a decision tree there is partial assignment that forces an input to follow this branch. As the branch is short we call it *consistent* if the corresponding assignment is consistent in the sense of Definition 3.1. In this paper it is always the case that all branches of a decision tree are consistent. This is achieved by simply erasing inconsistent branches.

We are interested in what happens to a decision tree T when subject to a (full) restriction σ or a partial assignment τ and the results are denoted by $T\lceil_\sigma$ and $T\lceil_\tau$, respectively. There is no essential difference between the two cases as in each case we have a decision tree where the values of some variables are already fixed and we just keep the paths consistent with these values.

Let us first state this in an operative manner. We start at the root of T and at each node we have a variable under consideration. If the value of this variables is forced by σ (or τ, respectively), the values along the path so far, and consistency, we choose the sub-tree with the consistent value and otherwise we explore both sub-trees in the natural way.

A more static way is to consider all paths of T from the root to a leaf and see which of the corresponding assignments are consistent with σ (or τ). The paths that are consistent remain and those not consistent are erased. It is easy to see that the remaining paths (possibly after some contractions) nicely fit into a decision tree and in fact the decision tree defined above.

When considering consistency with σ we of course make use of the information that all old variables that are governed by the same new variables must take equal or opposite values as governed by the negations of the new path variable x_P.

Lemma 5.1: If $n_1 + n_2 \leq o(n)$ where n is the size of the current grid, then if T is a decision tree of depth n_1 and τ is a partial assignment that gives values to at most n_2 variables then $T\lceil_\tau$ is a non-empty decision tree.

Proof: This follows from Lemma 3.2. ■

Lemma 5.2: Suppose σ is full restriction whose output is an instance of size n and let T be a decision tree of depth $o(n)$. Then $T\lceil_\sigma$ is a non-empty decision tree.

Proof: At each step going down a path on a decision tree at least one of the two values of a variable is consistent with σ and the path so far. ■

Lemma 5.3: If T is a decision tree of depth n_3 and let τ_1 and τ_2 are assignments that gives values to at most n_1 and n_2 variables, respectively, that are consistent with each other. Then, provided that $n_1 + n_2 + n_3 \leq o(n)$ we have $T\lceil_{\tau_1}\lceil_{\tau_2} = T\lceil_{\tau_2}\lceil_{\tau_1}$ and both are non-empty decision trees.

Proof: Taking the static view of restricted decision trees both contain all paths of T that are consistent with $\tau_1 \cup \tau_2$. ■

We let a *1-tree* be a decision tree where all leaves are labeled 1 and define a *0-tree* analogously. Special cases of such trees are trees of depth 0. Next we turn to a procedure of representing formulas by decision trees of small depth.

VI. t-EVALUATIONS

We have a supposed proof and we have the set of formulas that appear in the proof. We also have each sub-formula in each of these formulas and this gives a set of formulas Γ. We use t-evaluations φ, a concept introduced by [8], that map formulas to decision trees of depth at most t. Such mappings will not be total and we are interested in finding t-evaluations defined over as large set of formulas as possible. This is made possible by, at the same time as extending the range, applying a restriction. Let us define the desired properties required of t-evaluations.

1) The constant true is represented by a 1-tree and the constant 0 is represented by a 0-tree.
2) If F is an axiom of the Tseitin contradiction then $\varphi(F)$ is a 1-tree.
3) If $\varphi(F) = T$ then $\varphi(\neg F)$ is a decision tree with the same topology as T but where the value at each leaf is negated.
4) Suppose $F = \vee F_i$. Consider a leaf in $\varphi(F)$ and the assignment, τ leading to this leaf. If the leaf is labeled 0 then for each i $\varphi(F_i)\lceil_\tau$ is a 0-tree and if the leaf is labeled 1 then for some i, $\varphi(F_i)\lceil_\tau$ is a 1-tree.

The intuitive role of $\varphi(F)$ is that it represents the formula F as a function on all assignments that satisfy[2] "the relevant" local Tseitin constraints. Let us explicitly give the representation of the axioms and take $(x_{e_1} \vee x_{e_2} \vee x_{e_3} \vee x_{e_4})$ where e_i are the four edges incident to a center v. Naturally each variable is represented by a decision tree of depth one. This clause is represented by a decision tree of depth three with all leaves labeled 1 asking the variables x_{e_1}, x_{e_2}, and x_{e_3} in order. The only leaf that requires a little bit of thought to see that it is labeled 1 is the node where all three variables are zero. In this leaf x_{e_4} is reduced to a decision tree of depth 0 with label 1 as the only value of x_{e_4} consistent the three 0s is 1.

Note that we cannot represent this formula by a smaller tree as, by rule 4, for each 1-leaf, we must have an assignment that forces one of the decision trees for x_{e_i} to be a 1-tree.

Another example is the conjunction of all the axioms. As we do not have any $\wedge$-gates this is represented as the negation of the disjunction of the negations of all axioms. As we just saw, each axiom is represented by a 1-tree of depth 3 and hence its negation is a 0-tree of the same depth. Any disjunction of such trees can be represented by a decision tree of depth zero where the only leaf has label 0 and hence the representation of the negation of such a disjunction is a tree of depth 0 with label 1.

Thus we have constant one as a representation for a formula that, when interpreted in the natural way, evaluates to false on each input. The reason is that each sub-formula looks true in the local sense.

For a general set of formulas we cannot hope to have a t-evaluation for a small t and our plan is to proceed as follows for $i = 0, 1, 2 \ldots d$.

- We have a t-evaluation for all formulas of Γ that were originally of depth i.
- Pick a random restriction σ_i and extend the t-evaluation to all formulas of $\Gamma\lceil_{\sigma_i}$ of original depth at most $i+1$.

At the starting point, $i = 0$, each formula is a literal or a constant which is represented by a natural decision tree of depth at most 1 and we start by proving that t-evaluations are compatible with restrictions.

Lemma 6.1: Given a set of formulas Γ' and a t-evaluation φ whose range includes Γ' and let σ be a full restriction whose output is a grid of size n. Then, provided that $t < n$, $\varphi(F)\lceil_\sigma$ is a t-evaluation whose range includes $\Gamma'\lceil_\sigma$.

Proof: This is an easy consequence of the definition but let us go over the various possibilities. To start with, hitting

[2]This is achieved since we only consider paths in decision trees with are consistent.

a decision tree with a restriction can never increase the depth of the decision tree and hence all representations are decision trees of depth at most t. Note also that as $t < n$ the resulting tree is non-empty. We need to check the properties of a t-evaluation.

The first and second properties are obvious as a restriction does not change the fact that something is 1-tree or a 0-tree.

The third property is also rather obvious. The decision trees for F and $\neg F$ are effected the same way and there is nothing that can change that the corresponding leaves have labels that are the negations of each other.

For the fourth property consider any path in T that appears in $T\lceil_\sigma$ and the corresponding assignment τ which thus is consistent with σ. As already τ reduces the T_i in a good way, we need only observe that $T_i\lceil_\sigma\lceil_\tau$ is a non-empty decision tree and hence it is a 1-tree or a 0-tree as desired. ■

Now we eventually come to the key lemma of the entire argument.

Lemma 6.2: Let s' be an integer and $s = \max(s', t)$, then there is a constant A such that the following holds. Suppose there is a t-evaluation whose range includes $F_i, 1 \leq i \leq m$ and let $F = \vee_{i=1}^m F_i$. Let σ be a random restriction from the space of restrictions defined in Section IV. Then the probability that $F\lceil_\sigma$ cannot be represented by a decision tree of depth at most s' is at most

$$(As^{27}t\Delta^{-1})^{s'/27}.$$

We postpone the proof of this lemma to Section VII and see how to use it. We apply it with $s' = t = s = \frac{1}{2}n^{1/(58(d+1))}$ and $\Delta = s^{29}$ (and hence $T = 4s^{58}$).

We start with the original Tseitin contradiction on the $n \times n$ grid. Let $n_i = nT^{-i}$. We are going to choose a sequence of full restrictions σ_i mapping a grid of size n_i to a grid of size n_{i+1} randomly. Let σ_i^* be the composition of σ_0, σ_1, $\ldots \sigma_i$. As stated above, Γ is the set of sub-formulas that appear in an alleged proof and we let

$$\Gamma_i = \{F\lceil_{\sigma_i^*} \mid F \in \Gamma \wedge depth(F) \leq i\}.$$

Let f_i be the number of sub-formulas of depth at most i in Γ.

Lemma 6.3: With probability $1 - f_i(s/A)^{-s/27}$ there is a t-evaluation φ_i whose range includes Γ_i.

Proof: We prove the lemma by induction over i. For $i = 0$ we have the t-evaluation that maps each literal to its natural decision tree of depth 1 and constants to decision trees of depth 0.

When going from depth i to depth $i+1$ we need to define φ_{i+1} on all formulas originally of depth at most $i + 1$ and consider any such F.

1) For each F of depth i it is, by induction, in the range of φ_i and we can appeal to Lemma 6.1.
2) If F is of depth i then $\varphi_{i+1}(\neg F)$ is defined from $\varphi_{i+1}(F)$ negating the leaves.
3) For $F = \vee F_i$ where each F_i is of depth i we apply Lemma 6.2.

The only place where the extension might fail is under step three but, by Lemma 6.2, the probability of failure for any individual formula is at most $(s/A)^{-s/27}$ and as we have at most $f_i - f_{i-1}$ formulas of depth exactly i, the induction is complete. ■

As a final piece we establish that all formulas appearing in a short proof must be represented by 1-trees and as the constant false is represented by a 0-tree this is a contradiction. In order to prove this we must go over the derivation rules of our proof system. The details are not important and we choose the same rules as [11] and these are as follows.

- (Excluded middle) $(p \vee \neg p)$
- (Expansion rule) $(p \rightarrow p \vee q)$
- (Contraction rule) $(p \vee p) \rightarrow p$
- (Association rule) $p \vee (q \vee r) \rightarrow (p \vee q) \vee r$
- (Cut rule) $p \vee q, \neg p \vee r \rightarrow q \vee r$.

The below lemma is essentially a verification and the proof of it can be found in the full version of this paper.

Lemma 6.4: Suppose we have derivation using the above rules and using the Tseitin conditions in the $n \times n$ grid as axioms. Let Γ be the set of formulas appearing as sub-formulas of any formula in the given derivation and suppose that we have a t-evaluation whose range includes Γ where $t = o(n)$. Then each line in the derivation is mapped to a 1-tree. In particular we do not reach a contradiction.

We are now ready for the main theorem.

Theorem 6.5: Suppose that $d \leq \frac{\log n}{59 \log\log n}$, then, for sufficiently large n, any depth-d Frege refutation of the Tseitin contradiction on the $n \times n$ grid requires size $exp(\Omega(n^{1/58(d+1)}))$.

Proof: Suppose we have a refutation of size S and consider the corresponding set of sub-formulas Γ.

With the given choice of Δ we have $T \leq n^{1/(d+1)}$ and hence we have a $nT^{-d} \geq T$ sized grid remaining after σ_d^*. The probability that we fail to have t-evaluation of all formulas in Γ after σ_d^* is, by Lemma 6.2 bounded by $S(s/A)^{-s/27}$. The probability that we at any stage of the process we do not have between $.99Cs$ and $1.01Cs$ alive centers in a sub-square is bounded by $n^2e^{-\Omega(s)}$. As $s = \omega(\log n)$, the sum of these two failure probabilities, for sufficiently large n, is smaller than 1 and hence there exists a σ_d^* which makes all sub-formulas in the proof have a t-evaluation and such that the final restriction gives a grid of size at least T. As $t = o(T)$ we can appeal to Lemma 6.4 and the proof is complete. ■

We have an immediate corollary.

Corollary 6.6: Any polynomial size Frege refutation of the Tseitin contradiction on the $n \times n$ grid requires formulas of depth $\Omega(\frac{\log n}{\log\log n})$.

Finally we turn to the proof of the switching lemma which is the heart of the argument.

VII. Proof of the switching lemma

Let $T_i = \varphi(F_i)$. We create an *extended canonical* decision tree for $F\lceil_\sigma$ by going over the trees T_i one by one. If there is a path in T_i to a leaf with label 1 that is consistent with the information we have so far, we explore the variables of this path (and some extra variables). Let us proceed.

It is important that the constructed decision tree does not depend on the preferred values along the chosen paths but we may, and indeed we will, let it depend on other parameters and in particular we make use of the knowledge of the non-chosen centers.

As we go over the T_i's we have a set S of exposed centers and an information set I that, jointly with ρ, guides the construction of the decision tree.

For non-chosen centers in S we know their connected components in π and if one center in such a connected component belongs to S then so does the entire component. For chosen centers in S we have asked for the values of all remaining variables adjacent to these centers and this information is present in I. The one-answers are recorded in the form of a path while the zero answers as two non-edges supported at the two chosen centers that are the endpoints of the chosen path.

We go over the decision trees one by one and let us see what happens when we consider T_i. Take the first (in some fixed order) path in T_i that leads to a leaf labeled 1. For the variables appearing on this path we have unique values required to reach this leaf. We let the *forcing information*, J, be a set of edges and non-edges that, jointly with I and ρ, forces[3] all variables on this path, from now on called "the forceable path" to take these unique values. Furthermore we require.

1) If J contains a non-edge from a chosen center it also contains a non-edge in the "reverse direction". As an example if it contains a non-edge going left from chosen center v then J contains a non-edge going right from the chosen center in the sub-square to the left of v.
2) Neither I nor J contains a path between a chosen center and a non-chosen center.
3) The information sets I and J are jointly consistent with ρ and disjoint.

At any point in the above procedure, the information I comes from information in π and from queries in the decision tree with answers τ. Let us first see that the lack of forcing information implies that T_i is in fact reduced to a 0-tree.

Lemma 7.1: If there is no forcing information for T_i then $T_i\lceil_{\sigma\tau}$ is a 0-tree.

Proof: Suppose indeed that there is a path in T_i that leads to a 1-leaf and is consistent with σ and τ. This implies that we can extend τ to τ_1 such that we reach this leaf. In other words, σ and τ_1 jointly determines the value of each variable on this path.

We proceed to construct some forcing information J. Let us consider a variable x_e on the path. For e whose closest endpoint is not chosen we include the information from π on this closest endpoint in direction of e. If the closest endpoint of e is chosen then e may or may not be on the chosen path in its direction.

If e is on the chosen path then the information τ_1 must contain the value of the corresponding path-variable and we include that information in the form of an edge or two non-edges in J. If e is not on the chosen path then we choose some value to the path-variable in its direction from its closest endpoint that is consistent with τ_1 and choices for previous variable set in the current process. Given the value of this variable we include this in J either as an edge or two non-edges. We need to check that J is a valid forcing information.

The first property that it forces the input to follow the path is true by construction and we turn to the other properties needed.

As π only contains paths between two non-chosen centers and τ_1 and its extension only paths between two chosen center, we cannot have a path between a chosen and non-chosen center in J and we need to check consistency with I.

On the non-chosen centers, I contains some information from π and as the information in J on the non-chosen part is also from π this is consistent (any duplicated information can simply be dropped from J).

On the chosen centers we know that τ_1 is an extension of τ, the information obtained in the decision tree up to this point. As the information in I on the chosen centers is exactly given by τ and the information in J, which is consistent with τ_1, which is an extension of τ we conclude that J is consistent with I. As we did not give this forcing information in the construction of the extended canonical decision tree we can conclude that the assumed 1-path does not exist and the lemma follows. ■

If there is forcing information J we expose all centers in the support of J but also some additional centers as follows.

- For any non-chosen center v in the support of J we expose the centers in its connected component in π.
- We let the chosen centers in the support of J be the nodes supplied by the adversary in the matching game described in Section III. We apply Lemma 3.3 and expose also the partners of these nodes in the matching provided by P_M.

We note that if the support of the forcing set J is of size r then the number of exposed centers is at most $4r$ as we expose at most 3 more nodes for any non-chosen center and at most one extra node for any chosen center.

[3]Please remember, by Definition 4.3 for a variable to be forced we need to know the relevant information at its closest endpoint.

We now extend the information I by including the connected component from π of the non-chosen exposed centers. For the chosen centers we ask all variables adjacent to any exposed center. We record one-answers as an edge in I and zero-answers as two non-edges including the other endpoint of a potential chosen path, i.e. the chosen center in the adjacent sub-square in the given direction.

Given this extended I it is possible to tell whether the forceable path in T_i is traversed. This follows as for any variable on the path the closest endpoint is now exposed and for each exposed node we have information pieces in all four directions. If this path is indeed followed, the process is ended as $T_i\lceil_{\sigma\tau}$ is a 1-tree and the path of the decision tree can be terminated with label 1.

If the forceable path is not followed we continue the process by first looking at T_i under this new extended information I and searching for some new forcing information of a different 1-path and then looking at $T_{i'}$ for $i' > i$.

Finally, if all T_i's have been processed we terminate the path in the decision tree and label the leaf 0. This ends the description of the creation of the extended canonical decision tree for $F\lceil_\sigma$. We observe that we have created a decision tree that is a legitimate choice for $\varphi(F)$. Indeed at any leaf labeled 1 we have found a T_i that is reduced to a 1-tree and if all T_i have been processed then, by Lemma 7.1, this leaf in the decision tree is correctly labeled 0.

Note that this process depends on ρ and π but not, in a serious way, on the negations of the preferred values along the paths between the chosen centers. As we have no paths between chosen and non-chosen centers the only difference is that for variables on chosen paths in one case is forced by the path and in the other case by two non-edges and this does not cause any difference as the supports are identical. As this is of key importance let us record this as a lemma.

Lemma 7.2: Let σ_1 be obtained from ρ_1 and π and σ_2 from ρ_2 and π where ρ_1 and ρ_2 pick the same set of centers and fixed values. Assume furthermore that there for each chosen path P there is a bit c_P such that for each grid-edge e on P the preferred values of x_e differ by c_P in ρ_1 and ρ_2. Then the only difference between the extended canonical decision trees of $F\lceil_{\sigma_1}$ and $F\lceil_{\sigma_2}$ is the labeling of the internal edges.

In the decision tree, at round j, we ask all questions to a set of variables touching the chosen centers of the set S. We say that the answers are *closed* iff the answer to a query is one iff it corresponds to an edge in the dynamic matching created by P_M. The resulting information set is then closed in the already defined sense. The following lemma is now a consequence of Lemma 7.2.

Lemma 7.3: If the probability that $F\lceil_\sigma$ needs a decision tree of depth s' is at least q, then the probability that the extended canonical decision tree of $F\lceil_\sigma$ contains a closed path of length at least s' it at least $2^{-s'}q$.

In view of this lemma we complete the proof by analyzing the probability of such a closed path. This analysis is done using the labeling technique of Razborov [17]. In other words we take a ρ that contributes to the above event and create a ρ^* which is also a restriction but with fewer live centers. We then establish that given ρ^* and some extra information it is possible to reconstruct ρ. Noting that there are many fewer ρ^* than ρ and the extra information can be limited in size we get the desired conclusion. Thus we assume that we have such a closed path and we proceed to construct ρ^*.

For technical reasons we stop the creation of the extended canonical decision tree once we have exposed at least s' centers and we analyze the probability that we ever reach this point. Suppose this happens after the gth stage, where $g \leq s'$ as we expose at least one center in each stage.

At the end of the process we have a set, S^g, of exposed centers which is of cardinality at least s' and at most $s'+8t$, as we at each stage expose at most $8t$ centers. This follows as J contains at most $2t$ centers as the length of each path in T_{i_j} is at most t and we add at most 2 centers for each variable on the path. We later expose at most three more centers for each element in the support of J.

Let us look at the forcing information in stage j and introduce some notation. The forceable path appears in T_{i_j} and let J_j be the forcing information set. As we continue processing the same T_i after a stage is completed it might be the case that $T_{i_j} = T_{i_{j+1}}$, but then the forceable paths are different.

Consider any center v in the support J_j. It has information in some of its directions coming from I and J_j. If it has information in all four directions nothing needs to be done. Otherwise, take one direction for which the information is not known. If there are more directions in which there is no information, add a non-edge in any other such direction.

If we already have an odd number of edges next to v we add a non-edge in the final direction and otherwise we add an edge to a fresh center in the suitable sub-square. By a fresh center we mean a non-chosen center that is not an element of S^g and has not been used for an earlier J_j. As we use at most one fresh center for each element in S^g the number of non-fresh centers is at most $2|S_g| \leq 2s' + 16t$. As there are $.99Cs$ non-chosen centers in any sub-square there is always a fresh center to add provided that C is a large enough constant.

When we have processed all centers of J_j we have created a closed graph which extends the information set J_j and which we denote γ_j. This follows as for each even degree center we have added a fresh center that is of degree one. As discussed previously, closed graphs can be used to define restrictions with fewer live centers and we define ρ^* to be the restriction defined by ρ together with the graph $\gamma = \cup_{j=1}^{g}\gamma_j$. This is a standard restriction where all centers in the support of γ are now dead. We call these the *disappearing* centers.

Before we turn to the reconstruction process let us intro-

duce some notation for the information sets of the decision tree process. Let us see what happens at stage j.

On the non-chosen centers there is the information of some connected components of π, namely all the exposed centers and let $I_{j,n}$ denote the union of these components discovered in stage j. For the chosen centers the information is obtained by the decision tree. As the decision tree is closed this is given by a matching on the exposed chosen centers. On top of this we have the information of non-edges of non-exposed chosen centers in the direction of exposed chosen centers. Call this information on the chosen centers $I_{j,c}$ and let I_j be the union of $I_{j,n}$ and $I_{j,c}$. Furthermore let I_j^* denote $\cup_{i=1}^{j-1} I_i$, the information set gathered during the first $j-1$ rounds. It turns out to be convenient to consider $\cup_{i=j}^{g} \gamma_i$, the graphs added after stage j, and we let γ_j^* denote this graph.

The high level plan is now as follows. As γ_j extends the forcing information J_j we have that $(\rho, I_j^* \cup \gamma_j)$ and hence $(\rho, I_j^* \cup \gamma_j^*)$ forces the input to traverse the jth forceable path. This path should enable us to find a good fraction of the elements of γ_j as the closest endpoints of some variable(s) on this path. We then use some external information to find the rest of the elements of γ_j (as well as its graph structure). Finally we then use external information to reconstruct I_j and proceed with stage $j+1$.

As I_1^* is the empty set and $\gamma_1^* = \gamma$ the starting point of the decision process is (ρ, γ) which forces exactly the same variables as ρ^* and thus we know where to start. Although these two objects force the same variables the information content is different in that (ρ, γ) contains the information we are trying to recreate, the identity of the disappeared centers.

We let ρ_j^* be the restriction obtained from applying γ_j^* to ρ and at stage j we will be working with (ρ_j^*, I_j^*) instead of $(\rho, I_j^* \cup \gamma_j^*)$. Again these two objects force the same set of variables but have different information contents.

It is important to identify T_{i_j} and the forceable path but unfortunately it might not be the first 1-path traversed by (ρ_j^*, I_j^*). The reason for this is that we might reach a 1-leaf by a path using variables that would give forcing information that is not allowed. For instance when we make sure that γ_j is closed we add paths between chosen and non-chosen centers and this is not allowed as forcing information. Another more subtle problem is that of requiring the other endpoint of non-edges on chosen centers when used as forcing information. It turns out that it is difficult to make sure that the information at the other endpoint is consistent with the rest of the information.

Let I_j^{*-} be the information pieces of I_j^* with any piece supported on γ_j^* removed and let I_j^- be I_j with the same type of pieces taken away. The removed pieces are simple to describe.

Lemma 7.4: An information piece in I_j^* that is on a center in the support of γ_j^* is in the form of a non-edge from a chosen center in the direction of an exposed chosen center.

Proof: The information set I_j^* consists of a closed graph jointly with non-edge information on chosen centers of the type allowed in the lemma. Since any information set J_i for $i \geq j$ is disjoint with I_j^* no γ_i with $i \geq j$ can intersect the closed graph part of I_j^*. ■

We get a direct consequence of Lemma 7.4.

Lemma 7.5: Any variable forced by (ρ, I_j^*) is forced also by (ρ_j^*, I_j^{*-}).

Proof: The removed pieces of I_j^* are, by Lemma 7.4, on centers that have disappeared in ρ_j^* and hence any variable forced by such a piece is fixed in ρ_j^*. As the piece of information is a non-edge in both I_j^* and γ_j^* it is forced to the same value. ■

As stated above we might have some 1-path before the forceable path of stage j. This is, in some vague sense be good, in that it reveals some element of γ. As we cannot count on this happening, however, this possibility is only a problem and we have to be careful to make sure that the reconstruction process is not fooled. Towards this end we introduce the *signature* of any disappearing center, v, as follows.

1) The value of j such that $v \in \gamma_j$.
2) The information of whether it is a closest endpoint to any variable on the forceable path and in such a case in which direction(s) it has variables appearing on this path.

Let us now describe the reconstruction procedure formally. It has the following information.

1) A counter j of the current stage to be reconstructed. Initially $j = 1$.
2) The restriction ρ_j^*. Initially $\rho_1^* = \rho^*$ and we describe below how to update.
3) The information set I_j^{*-}. Initially this is empty and we describe below how to update.
4) A set E of disappearing centers together with their signatures. Initially E is empty.

In the reconstruction process we need to find the identity of some centers. Let us discuss different contexts where this happens and how much external information is needed. For some disappearing centers we also specify the signature which amounts to $O(s)$ possibilities for each center. We have the following cases.

1) A disappearing center that is the closest endpoint of a variables on a discovered 1-path. This can be found by giving the distance from the root on the path at cost t.
2) A disappearing center that is not the closest endpoint of a variable on a path but we know the sub-square where it is located. This can be specified at cost Δ.
3) A non-disappearing and live center where we know the sub-square. This can be specified at cost $1.01Cs$ as these are the number of live centers in any sub-square.

The two first situations appear when finding centers in γ_j while the last situation appears when finding centers in I_j. Identifying a disappearing center has "profit" (as will be seen in the final calculation of counting the number of ρ^* compared to the number of ρ) of $\Omega(\Delta/s)$ and thus there is a huge profit in the first case and the moderate loss in the second. For the third case there is no associated profit but on other hand only a moderate cost. The key for the final analysis is to bound the number of costly step by a constant times the number of profitable steps of the first kind. Let us now formally define the reconstruction process.

1) Find the next 1-path traversed by (ρ_j^*, I_j^{*-}).
2) Locate the closest endpoints of all variables on this path. If any such center belongs to E and its signature does not match the current path, go to the next 1-path. By "not matching" we mean that the stage information is incorrect or that the direction(s) of the edges involved does not exactly match the signature.
3) Read a bit b to determine if there are more disappearing centers to be found as the closest endpoint to variables on this path.
4) If $b = 1$ read one integer that is at most t to determine a disappearing center that is the closest endpoint of a variable on this path. Read its signature. If this signature agrees with the current path repeat step 3 and otherwise include it in E and go to the next path.
5) If $b = 0$ we have found the forceable path. We read some external information to determine γ_j and I_j^- (details below). Update ρ_j^* to ρ_{j+1}^* and I_j^{*-} to I_{j+1}^{*-}, drop any disappearing center of stage j from E, increase j and repeat from 1.

The are a few details and facts about this reconstruction procedure to sort out. Let us start with establishing that we are indeed correctly identifying the forceable path.

Lemma 7.6: If a 1-path is the first path to be forced by (ρ_j^*, I_j^{*-}) and the signatures of all closest endpoints of all variables on this path match, then this path is the jth forceable path.

Proof: As all variables on the path are forced we must have the information of their closest endpoints in the correct direction. As none of the variables have a closest endpoint of a stage later than j, and the signatures are correct, the path is forced by $(\rho, I_j^{*-} \cup J_j)$ jointly possibly with a non-edge in γ_j contained in I_j^*. This implies that the forcing information J_j is valid for this path and being the first such path it must be the jth forceable path. ■

Let us now see how to reconstruct γ_j. We have already identified all the closest endpoints of variables on the forceable path and we know, by their signature which directions they need a neighbor. We read the identity of these centers at a cost[4] of at most Δ for each center. This identifies J_j. To finalize the description of γ_j we read the identity of the fresh centers used to make γ_j closed at a cost of Δ for each such center. Having identified γ_j we turn to I_j^-.

We have a bit for each element in γ_j to indicate whether it is also an element of I_j and we proceed to identify the rest of I_j. We first reconstruct the missing non-chosen centers. For each non-chosen center in J_j using $O(1)$ bits we first find out the size of the connected component in π and the directions of each edge. Then we identify each such center at cost $1.01Cs$. Here we use the fact that as these variables are part of I_j they cannot be included in the support γ_{j+1}^* and hence they are alive in ρ_j^*.

For the chosen centers we can again discover the graph part with $O(1)$ bits per center for structure and an integer of size $1.01Cs$ for the identity (as also these are alive in ρ^*. The non-edges not supported on γ_j^* are also reconstructed at cost $1.01Cs$ for identity and $O(1)$ bits per center for direction.

Finally for any center in γ_j we have 4 bits to describe whether the piece of information in the form of non-edge in any direction(s) should be added in I_{j+1}^{*-}.

This terminates the description of the reconstruction and let us sum up the external information needed. Let a_j be the number of disappearing centers that are discovered through being the closest endpoint of a discovered variable and are part of the jth forceable path and let b_j the number of additional centers in γ_j. Furthermore let c_j the the number of centers needed to be discovered in I_j^- after γ_j was discovered.

Lemma 7.7: We have $b_j + c_j \leq 25a_j$.

The fact that there is some constant such that the above lemma is true is, hopefully, quite believable but getting the best constant requires some case analysis. We leave the proof of this lemma for the full version of the paper.

Now we are ready to make the final calculation. Letting $a = \sum_{j=1}^{g} a_j$ and defining b and c similarly we can add up the extra information as follows.

- The disappearing centers that are discovered as closest endpoints contribute a factor t^a.
- The other disappearing centers a contribute a factor at most Δ^b (or less as discussed in the footnote).
- The signatures contribute at most $(As')^a$ for a constant A as signatures are only needed for disappearing centers discovered as closest endpoints.
- The centers discovered to be part of I contribute a factor $(1.01Cs)^c$.
- The graph structure of γ and I as well as the information which elements of γ_j are included in I_j contributes a factor B^{a+b+c}, for some constant B.
- The bits b contribute $2^{s'+8t+s'}$. This follows as we can have at most $s' + 8t$ bits that are 1 (as each time a

[4]It might be the case that some of these centers are uniquely determined and/or found previously and are part of E. In such a case the cost, including the signature is $O(st)$ which is much lower.

disappearing variable is discovered) and at most s' bits that are 0 (as each time a stage is ended).

Let $m = \Delta(n/T)^2$ be the total number of centers. The number of ways to choose ρ^* is $2^{r_1}\binom{m}{k-(b+a)}$ where 2^{r_1} is the number of possibilities for the choice of fixed and preferred variables once the choice of centers is fixed. Similarly the number of choices for ρ is $2^{r_1}\binom{m}{k}$. This implies that the probability of having a described closed path is bounded by

$$\frac{t^a \Delta^b s^a s^c A^{a+b+c} 2^{r_1}\binom{m}{k-(a+b)}}{2^{r_1}\binom{m}{k}} \tag{1}$$

for some (modified) absolute constant A. The quotient of the the binomial coefficients equals

$$\prod_{i=0}^{a+b-1} \frac{k-i}{m+i-k} \leq \left(\frac{k}{m-k}\right)^{a+b} = \left(\frac{Cs}{\Delta - Cs}\right)^{a+b} \tag{2}$$

$$\leq \Delta^{-(a+b)} s^{a+b} A^{a+b}, \tag{3}$$

for some (again different) constant A. We conclude that the probability of the closed path in the decision tree we are analyzing is at most

$$\Delta^{-a} s^{2a+b+c} t^a A^{a+b+c}, \tag{4}$$

for again a new constant A. Applying Lemma 7.7 and modifying A we have that this is bounded by

$$\Delta^{-a} s^{27a} t^a A^a = (As^{27} t \Delta^{-1})^a. \tag{5}$$

Finally as the number of exposed centers is at most $a + b + c$ we have $a + b + c \geq s'$ and hence $a \geq s'/27$ and this concludes that analysis of the probability of a closed path. Lemma 6.2 now follows from Lemma 7.3 and a final modification of the constant A.

ACKNOWLEDGEMENT

Some early ideas of this paper were discussed with Pavel Pudlak and Jakob Nordström. I am also grateful for later discussions with Ilario Bonacina, Susanna F. de Rezende Marc Vinyals, Joseph Swernofsky and Mladen Mikša.

REFERENCES

[1] G. S. Tseitin, "On the complexity of derivation in the propositional calculus," in *Studies in constructive mathematics and mathematical logic, Part II*, A. O. Slisenko, Ed., 1968.

[2] A. Haken, "The intractability of resolution," *Theoretical Computer Science*, vol. 39, pp. 297 – 308, 1985.

[3] E. Ben-Sasson and A. Wigderson, "Short proofs are narrow–resolution made simple," *Journal of the ACM*, vol. 48, no. 2, pp. 149–169, 2001.

[4] M. Ajtai, "The complexity of the pigeonhole principle," *Combinatorica*, vol. 14, no. 4, pp. 417–433, 1994.

[5] S. Bellantoni, T., and A. Urquhart, "Approximation and small-depth frege proofs," *SIAM J. Comput.*, vol. 21, no. 6, pp. 1161–1179, 1992.

[6] T. Pitassi, P. Beame, and R. Impagliazzo, "Exponential lower bounds for the pigeonhole principle," *Computational Complexity*, vol. 3, pp. 97–140, 1993.

[7] J. Krajícek, P. Pudlák, and A. R. Woods, "An exponenetioal lower bound to the size of bounded depth frege proofs of the pigeonhole principle," *Random Struct. Algorithms*, vol. 7, no. 1, pp. 15–40, 1995.

[8] A. Urquhart and X. Fu, "Simplified lower bounds for propositional proofs," *Notre Dame Journal of Formal Logic*, vol. 37, no. 4, pp. 523–544, 1996.

[9] E. Ben-Sasson, "Hard examples for the bounded depth frege proof system," *Computational Complexity*, vol. 11, no. 3-4, pp. 109–136, 2002.

[10] S. Buss, "Polynomial size proofs of the propositional pigeonhole principle," *Journal of Symbolic Logic*, vol. 52, pp. 916–927, 1987.

[11] T. Pitassi, B. Rossman, R. A. Servedio, and L.-Y. Tan, "Poly-logarithmic frege depth lower bounds via an expander switching lemma," in *Proceedings of the Forty-eighth Annual ACM Symposium on Theory of Computing*, ser. STOC '16. New York, NY, USA: ACM, 2016, pp. 644–657.

[12] M. Furst, J. Saxe, and M. Sipser, "Parity, circuits and the polynomial-time hierarchy," *Mathematical Systems Theory*, vol. 17, pp. 13–27, 1984.

[13] M. Sipser, "Borel sets and circuit complexity," in *Proceedings of the fifteenth annual ACM symposium on Theory of computing*, ser. STOC '83. New York, NY, USA: ACM, 1983, pp. 61–69.

[14] A. C.-C. Yao, "Separating the polynomial-time hierarchy by oracles," in *Foundations of Computer Science, 1985., 26th Annual Symposium on*, oct. 1985, pp. 1 –10.

[15] J. Håstad, "Almost optimal lower bounds for small depth circuits," in *Proceedings of the eighteenth annual ACM symposium on Theory of computing*, ser. STOC '86. New York, NY, USA: ACM, 1986, pp. 6–20.

[16] J. Mehta, "Tree tribes and lower bounds for switching lemmas," *CoRR*, vol. abs/1703.00043, 2017. [Online]. Available: http://arxiv.org/abs/1703.00043

[17] A. A. Razborov, *Bounded Arithmetic and Lower Bounds in Boolean Complexity*. Boston, MA: Birkhäuser Boston, 1995, pp. 344–386, editors Peter Clote and Jeffrey Remmel.