On Semi-Algebraic Proofs and Algorithms

Noah Fleming  
University of California, San Diego, CA, USA  
Memorial University, St. John’s, Canada

Mika Göös  
EPFL, Lausanne, Switzerland

Stefan Grosser  
McGill University, Montreal, Canada

Robert Robere  
McGill University, Montreal, Canada

Abstract

We give a new characterization of the Sherali-Adams proof system, showing that there is a degree-$d$ Sherali-Adams refutation of an unsatisfiable CNF formula $C$ if and only if there is an $\varepsilon > 0$ and a degree-$d$ conical junta $J$ such that $\text{viol}_C(x) - \varepsilon = J$, where $\text{viol}_C(x)$ counts the number of falsified clauses of $C$ on an input $x$. Using this result we show that the linear separation complexity, a complexity measure recently studied by Hrubeš (and independently by de Oliveira Oliveira and Pudlák under the name of weak monotone linear programming gates), monotone feasibly interpolates Sherali-Adams proofs.

We then investigate separation results for $\text{viol}_C(x) - \varepsilon$. In particular, we give a family of unsatisfiable CNF formulas $C$ which have polynomial-size and small-width resolution proofs, but for which any representation of $\text{viol}_C(x) - 1$ by a conical junta requires degree $\Omega(n)$; this resolves an open question of Filmus, Mahajan, Sood, and Vinyals. Since Sherali-Adams can simulate resolution, this separates the non-negative degree of $\text{viol}_C(x) - 1$ and $\text{viol}_C(x) - \varepsilon$ for arbitrarily small $\varepsilon > 0$. Finally, by applying lifting theorems, we translate this lower bound into new separation results between extension complexity and monotone circuit complexity.

2012 ACM Subject Classification  Theory of computation → Proof complexity

Keywords and phrases  Proof Complexity, Extended Formulations, Circuit Complexity, Sherali-Adams

Digital Object Identifier  10.4230/LIPIcs.ITCS.2022.69

Funding  Noah Fleming: Research supported by NSERC. Robert Robere: Research supported by NSERC.

1 Introduction

The Sherali-Adams hierarchy is a well-studied method in optimization that provides an automatic way to convert a linear program with “weak” approximation guarantees into a linear program with “strong” approximation guarantees. Each “level” of the Sherali-Adams hierarchy is defined by systematically adding new variables and inequalities to the original linear program – “lifting” it to a higher dimensional space – and then “projecting” it back down to a polytope contained within the original linear program. The Sherali-Adams hierarchy at level $r$ results in a linear program with roughly $(n r)$ constraints and variables, and it is known that at level $n$ the hierarchy converges to the integral hull of the starting polytope. Owing to its strength and its generality, much work has been spent on understanding the strength of the Sherali-Adams hierarchy and its subsystems when applied to NP-Hard optimization problems [1, 6, 15, 19, 38, 43, 45, 46, 49, 56], as well as its strength when it is treated as a refutation system in propositional proof complexity [2, 5, 7, 18, 21, 34, 36].
A powerful generalization of the Sherali-Adams hierarchy is the notion of an extended formulation, which was originally formulated by Yannakakis [58] and extended by Braun et al. [9] to nested pairs of polytopes. Given a pair of polytopes $P \subseteq Q \subseteq \mathbb{R}^n$, an extended formulation of the pair $(P, Q)$ is a polytope $K \subseteq \mathbb{R}^m$ with $m \geq n$ along with a projection $\pi$ such that $P \subseteq \pi(K) \subseteq Q$. Indeed, the Sherali-Adams hierarchy gives many examples of extended formulations: if we are given some linear program $Q \subseteq [0,1]^n$ which gives a “weak” approximation guarantee to any point in the integral hull

$$\text{int}(Q) := \text{conv}\{x \in \mathbb{Z}^n \mid x \in Q\},$$

then each level of the Sherali-Adams hierarchy produces an extended formulation $(K, \pi)$ such that

$$\text{int}(Q) \subseteq \pi(K) \subseteq Q.$$

We regard this as giving a “tighter” relaxation to the integral hull, when compared to the starting polytope $Q$ (indeed, at $n$-th level of the hierarchy it turns out that $\pi(K) = \text{int}(Q)$).

More generally, for some polytopes $P \subseteq \mathbb{R}^n$ with exponentially many facets, it turns out to be possible to find an extended formulation $K \subseteq \mathbb{R}^m$ such that $\pi(K) = P$ but $m = \text{poly}(n)$ and $K$ has only polynomially many facets [58] (this is certainly a boon if $P = \text{int}(Q)$ for some combinatorial optimization problem encoded by $Q$!). Given a polytope pair $P \subseteq Q$ we therefore let $\chi(P, Q)$ denote the size (= number of facets) of the smallest extended formulation for the pair $P, Q$. After a breakthrough result by Fiorini et. al. [25], strong lower bounds have been shown for the extension complexity of polytopes associated with many standard NP-Hard optimization problems [9–11, 14, 25, 31, 42, 54, 55, 58]. All of these lower bound results crucially rely on the close relationship between the number of facets in any extended formulation and the non-negative rank of a certain related matrix [9, 58].

**Theorem 1 (Factorization Theorem).** Let $P \subseteq Q \subseteq \mathbb{R}^n$ be polyhedral sets, let $v_1, \ldots, v_n$ be the vertices of $P$ and let $a_1 \cdot x \leq b_1, \ldots, a_m \cdot x \leq b_m$ be linear inequalities in $\mathbb{R}^n$ describing the facets of $Q$. The size of the smallest extended formulation of $(P, Q)$ is $\text{rank}^+(S_{P, Q}) + 1$ where $S_{P, Q}$ is the $n \times m$ matrix defined by $S_{P, Q}(i, j) = b_j - a_j \cdot v_i$.

Indeed, works of Chan et. al. [14], Göös et. al. [33], and Kothari, Raghavendra, and Steurer [42] have shown that for certain NP-Hard optimization problems, lower bounds on the size of arbitrary extended formulations follow immediately from lower bounds on the Sherali-Adams hierarchy. These results join a long line of lifting theorems in communication, proof, and circuit complexity [16, 17, 22, 23, 26, 29, 30, 32, 35, 50, 52, 57] which systematically relate the complexity of computations in “complicated” computational models with complexity in “simple” computational models.

### 1.1 Sherali-Adams as Proofs and Extended Formulations as Circuits

In this work we further the study of the Sherali-Adams hierarchy and its relationship with extended formulations, but approach it from a different point of view. In particular, we are interested in the “dual view” of Sherali-Adams as a propositional proof system (first considered in [18]), and in extended formulations as a device for computing boolean functions. Before we formally state our results, let us first describe these two perspectives.
A conical junta $\mathcal{J}$ is a non-negative linear combination of conjunctions over a set of $\{0,1\}$-valued variables. Given an unsatisfiable CNF formula $\mathcal{C} = C_1 \land C_2 \land \cdots \land C_m$, a Sherali-Adams refutation\(^1\) of $\mathcal{C}$ is given by a list of $m + 1$ conical juntas $\mathcal{J}_1, \ldots, \mathcal{J}_{m+1}$ such that

$$
\sum_{i=1}^{m-1} -\mathcal{C}_i \mathcal{J}_i + \mathcal{J}_{m+1} = -1
$$

where all operations are done in multilinear polynomial arithmetic (so, $x^2 = x$) over $\mathbb{R}$ and $\overline{C}_i$ is the negation of the clause $C_i$. Indeed, if such a list of conical juntas exists then the original CNF formula must indeed be unsatisfiable, as if $x$ was a satisfying assignment then $\overline{C}_i(x) = 0$ for all $C_i$ and the above expression would reduce to $-1 = \mathcal{J}_{m+1}(x) \geq 0$, a contradiction. In this way, Sherali-Adams is naturally viewed as a proof system for refuting unsatisfiable formulas, and we can discuss complexity measures of its proofs such as the degree (i.e. the maximum degree of any product $\overline{C}_i \mathcal{J}_i$ as a multilinear polynomial) and its size\(^2\) (the number of distinct monomials occurring in the proof after expanding all products and before cancellations).

It is also quite natural to study extended formulations as non-uniform computation devices for boolean functions. The next definition was formally introduced by Hrubeš [40].

\begin{definition}
A separating polytope for $f : \{0,1\}^n \to \{0,1\}$ is a polytope $P \subseteq \mathbb{R}^n$ such that $\text{conv } f^{-1}(1) \subseteq P$ and $P \cap f^{-1}(0) = \emptyset$. We say a polyhedron $P \subseteq \mathbb{R}^n$ is monotone if $x \in P \Rightarrow y \in P$ whenever $x \leq y$, and if $P \subseteq \mathbb{R}^n$ we let $P^* := \{ y \in \mathbb{R}^n : \exists x \in P : x \leq y \} \supseteq P$ be the monotone closure of $P$. A monotone separating polytope for $f$ is a polytope $P$ such that $P^*$ is a separating polytope for $f$.
\end{definition}

Observe that if we have such a separating polytope $K$, then when given $x \in \{0,1\}^n$ we can test if $f(x) = 1$ simply by testing if $x \in K$. One natural way to construct such separating polytope is as follows, and was introduced by Hrubeš [39] and studied independently by Göös, Jain and Watson [31]. Given a boolean function $f : \{0,1\}^n \to \{0,1\}$ define the polytope $Q_{f,1}$ by the $|f^{-1}(0)|$ linear inequalities

$$
\forall y \in f^{-1}(0) : \sum_{i=1}^{n} x_i (1 - y_i) + (1 - x_i)y_i \geq 1
$$

where we note that $\sum_i x_i (1 - y_i) + (1 - x_i)y_i =: h(x,y)$ is exactly the hamming distance between two $\{0,1\}$-valued vectors $x, y$. It is easy to see that $x \in Q_{f,1}$ for every $x \in f^{-1}(1)$ since $x$ must differ from every $y \in f^{-1}(0)$ on some coordinate; similarly, for any $y \in f^{-1}(0)$ we have that $y \notin Q_{f,1}$ as the hamming distance from $y$ to itself is $0$. It therefore follows that any extended formulation of the pair $(\text{conv } f^{-1}(1), Q_{f,1})$ yields a separating polytope as we described above. One can further specialize this construction when the function $f : \{0,1\}^n \to \{0,1\}$ is monotone (recall $f$ is monotone if $x \leq y \Rightarrow f(x) \leq f(y)$, where the first inequality is interpreted coordinate-wise). In this case, we can simplify the description

\footnote{This encoding of Sherali-Adams is slightly different from the “usual” definition of Sherali-Adams as a refutation system of linear inequalities or polynomial equations, but is easily seen to be equivalent (cf. Claim 3.32 in [27]).}

\footnote{This measure is sometimes called the monomial size in other works on the Sherali-Adams hierarchy to differentiate from the bit-length of the encoding of the proof. As this is the natural notion of size for our purposes we simply call it size.}
Figure 1 A monotone separating polytope.

of $Q_{f,1}$ and still obtain a separating polytope. Specifically, if $f$ is monotone then define the polytope $Q^+_{f,1}$ by the linear inequalities

$$\forall y \in f^{-1}(0) : \sum_{i=1}^n x_i(1 - y_i) \geq 1.$$ 

The fact that an extended formulation of $(\text{conv } f^{-1}(1), Q^+_{f,1})$ is a separating polytope (indeed, now we can even have a monotone separating polytopes) for monotone $f$ follows by a similar argument as before, but uses the additional fact that for monotone functions, $f(x) = 1$ and $f(y) = 0$ if and only if there is a coordinate $i$ such that $x_i = 1, y_i = 0$. We depict such a pair in Figure 1.

By using the Factorization Theorem (Theorem 1), we can relate the size of these extended formulations to the non-negative ranks of certain matrices. Namely, given $f : \{0,1\}^n \rightarrow \{0,1\}$ define the $f^{-1}(1) \times f^{-1}(0)$ matrix $S_f(x, y) := \sum_{i=1}^n x_i(1 - y_i) + (1 - x_i)y_i$, and its “monotone” counterpart $S^+_f(x, y) := \sum_{i=1}^n x_i(1 - y_i)$. Define the separation complexity quantities

$$\text{sep}_1(f) := \text{rank}^+(S_f - 1), \quad \text{msep}_1(f) := \text{rank}^+(S^+_f - 1)$$

where $\mathbf{1}$ represents the all-1s matrix of the appropriate complexity. By Theorem 1, these two quantities capture the size of the smallest extended formulations for the pairs $(\text{conv } f^{-1}(1), Q_{f,1})$ and $(\text{conv } f^{-1}(1), Q^+_{f,1})$, respectively.

Now, it is natural to ask: are these interesting complexity measures of boolean computation? Hrubeš [39] and Göös, Jain and Watson [31] showed that the answer is yes: both of these quantities yield lower bounds on the formula complexity of computing $f$!

**Theorem 3.** For any boolean function $f : \{0,1\}^n \rightarrow \{0,1\}$, we have $\text{sep}_1(f) = O(n \cdot F(f))$ where $F(f)$ is the size of the smallest boolean formula computing $f$. Furthermore, if $f$ is monotone, then $\text{msep}_1(f) = O(n \cdot mF(f))$ where $mF(f)$ is the size of the smallest monotone boolean formula computing $f$.

Furthermore, observe the following. If our goal is just to separate the 1s of $f$ from the 0s of $f$ by means of a polytope we can weaken the inequalities

$$\sum_{i=1}^n x_i(1 - y_i) + (1 - x_i)y_i \geq 1, \quad \forall y \in f^{-1}(0).$$
conv \((f^{-1}(1))\)

**Figure 2** The left depicts the polytope pair \(\text{conv } f^{-1}(1)\) and \(Q_{f,1}\); the right depicts \(\text{conv } f^{-1}(1)\) and \(Q_{f,\varepsilon}\) for some \(\varepsilon < 1\). By setting \(\varepsilon < 1\) it is potentially easier to project a relatively low-facet polytope separating the 0s from the 1s of \(f\).

Since \(x \in f^{-1}(1)\) and \(y \in f^{-1}(0)\) are boolean strings, by integrality it suffices to have the inequalities

\[
\forall y \in f^{-1}(0) : \sum_{i=1}^{n} x_i (1 - y_i) + (1 - x_i) y_i \geq \varepsilon
\]

for any \(\varepsilon > 0\) (note we must have \(\varepsilon > 0\), since if \(\varepsilon = 0\) then \(y \in f^{-1}(0)\) will occur inside of the polytope). With this in mind, define the quantities

\[
\text{sep}(f) := \min_{\varepsilon > 0} \text{rank}^+(S_f - \varepsilon I), \quad \text{msep}(f) := \min_{\varepsilon > 0} \text{rank}^+(S_f^+ - \varepsilon I),
\]

which generalize \(\text{sep}_1(f)\) and \(\text{msep}_1(f)\) in the natural way. These quantities were recently studied by Hrubeš [40], where it was shown that they lower bound the size of boolean circuits computing \(f\).

**Theorem 4.** For any boolean function \(f : \{0, 1\}^n \rightarrow \{0, 1\}\) we have \(\text{sep}(f) = O(C(f) + n)\), where \(C(f)\) is the size of the smallest boolean circuit computing \(f\). Furthermore, if \(f\) is monotone, then \(\text{msep}(f) = O(mC(f) + n)\), where \(mC(f)\) is the size of the smallest monotone boolean circuit computing \(f\).

We also note that the quantity \(\text{msep}(f)\) was independently studied by Pudlák and de Oliveira Oliveira, where it was captured by a model of computation they called weak monotone linear programming gates [20].

To summarize: the Sherali-Adams hierarchy can be interpreted both as a particular family of extended formulations in optimization, and also as a natural family of proof systems for refuting unsatisfiable formulas. If we consider extended formulations more generally, then it is natural to interpret them as a computational devices whose complexities are related to standard boolean circuit and formula models. In this way, the known “lifting theorems” from Sherali-Adams proofs to extended formulations fall naturally in line with other proof-to-circuit lifting theorems [22, 29, 52], as we will show next.
1.2 Our Results

Normal Form for Sherali-Adams

In this work, we systematically relate the complexity of Sherali-Adams proofs and its fragments with the linear separation quantities $\text{sep}(f)$, $\text{msep}(f)$, $\text{sep}(X)$, and $\text{msep}(X)$. Our first main result is the following novel “normal form” for Sherali-Adams proofs. Given a CNF formula $F = C_1 \land C_2 \land \cdots \land C_m$ let $\text{viol}_C(x) = \sum_{i=1}^{m} C_i(x)$ denote the number of falsified clauses on an input $x$. Furthermore, if $f : \{0, 1\}^n \rightarrow \mathbb{R}_{\geq 0}$ is a non-negative real-valued boolean function then let $\text{deg}^+(f)$ denote the minimum degree of a conical junta $J$ such that $f = J$.

**Theorem 5.** For any unsatisfiable CNF formula $C$, if there is a Sherali-Adams refutation of $C$ with degree $d$ and size $s$, then there is an $\epsilon > 0$ and a degree $d$, size $s2^d$ conical junta $J$ such that $\text{viol}_C - \epsilon = J$.

Consequently, if $\text{SA}(C)$ is the minimum degree of any Sherali-Adams refutation of $C$, then

$$\text{SA}(C) = \min_{\epsilon > 0} \text{deg}^+(\text{viol}_C - \epsilon).$$

Several remarks on this theorem are in order. First, although we have stated this theorem for CNF formulas, a similar result immediately follows for arbitrary boolean CSPs. This is because if $P : \{0, 1\}^k \rightarrow \{0, 1\}$ is an arbitrary boolean CSP then we can represent $P$ as a width-$k$ unique DNF of its 1-inputs. This means that $\neg P = \sum_i D_i$ for some conjunctions $D_i$ (specifically, conjunctions that recognize the 0s of $P$). By substituting these sums for the predicates we can immediately deduce a more general theorem for refuting arbitrary boolean CSPs.

Second, if there is a conical junta $J$ such that $\text{viol}_C - \epsilon = J$ then note that we immediately obtain a Sherali-Adams refutation by rearranging the expression:

$$\sum_{i=1}^{m} \frac{-C_i}{\epsilon} + \frac{J}{\epsilon} = -1.$$

Thus this truly is a normal form that preserves the degree and the size (although, the size is only preserved up to a $2^d$ factor).

Third, one should note the similarities between Theorem 5 and the definitions of $\text{sep}(f)$ and $\text{msep}(f)$ – they are identical up to the substitution of $\text{deg}^+$ for $\text{rank}^+$ and $\text{viol}_C$ for $S_f/S_f^+$. Looking forward, this similarity turns out to be quite important for the rest of our results.

Finally, we remark that the fragment of Sherali-Adams corresponding to $\text{deg}^+(\text{viol}_C - 1)$ has also been studied in the literature as an object of interest. Göös, Jain and Watson studied it under the name of the “$3-1$ Game”, in which they proved degree lower bounds for Tseitin formulas [31]. Filmus, Mahajan, Sood, and Vinyals studied a further restriction in which the coefficients are required to be integers; they showed that this restriction was closely related to the complexity of MaxSAT resolution [24].

Feasible Interpolation for Sherali-Adams

Next, inspired by the normal-form theorem, we further develop the connection between Sherali-Adams and separation complexity by way of *monotone feasible interpolation*. 
The feasible interpolation method relates the complexity of proofs to the complexity of computational models. Suppose we are given an unsatisfiable CNF formula of the form $A(x,z) \land B(y,z)$. Then, under an assignment $z \mapsto \alpha$, at least one of the formulas $A(x,\alpha)$, $B(y,\alpha)$ is unsatisfiable. Thus, we can associate with $A \land B$ a partial function $I : \{0,1\}^z \to \{0,1\}$, known as an interpolant, satisfying

$$I(\alpha) = \begin{cases} 0 & \text{if } A(x,\alpha) \text{ is satisfiable} \\ 1 & \text{if } B(y,\alpha) \text{ is satisfiable} \end{cases}$$

In its original form, introduced at this level of generality in the classic work of Krajiček [44], we say that a proof system $P$ has feasible interpolation if we can extract a small computation of the interpolant $I$ in some computational model from any small $P$-proof of $A \land B$. Furthermore, if $A \land B$ satisfies a certain monotonicity property, namely that all $z$-literals occur only negatively in $A$, then the interpolant function $I$ is monotone and we say that $P$ has monotone feasible interpolation if we can extract from a $P$-proof of $A \land B$ a computation in some monotone computational model. Instantiations of the method of (monotone) feasible interpolation have led to a number of important lower bounds in proof complexity. Razborov [51] showed that from proofs in certain fragments of bounded arithmetic one could extract Boolean circuits and used this to establish conditional unprovability of $P \neq \text{NP}$ in these systems. The first lower bounds for Cutting Planes were established by Pudlák [47], who showed that proofs in this system gave rise to monotone real circuits; this built upon earlier work by Bonet, Pitassi and Raz [8] who showed that low-weight Cutting Planes gave rise to monotone circuits. Pudlák showed that span programs monotone feasibly interpolate Nullstellensatz [48], and de Oliveira Olivera and Pudlák showed that proofs in the Lovász-Schrijver system convert to monotone linear programming circuits [20].

Recently, Hrubeš and Pudlák [41] and Fleming et al. [28] showed that the method of monotone feasible interpolation could be generalized to work for arbitrary unsatisfiable CNF formulas, not only for split formulas. They showed that from a small Cutting Planes proof of an unsatisfiable CNF formula $C$ one could extract small monotone circuit computing an associated monotone function, termed the unsatisfiability certificate by Hrubeš and Pudlák [41]. (Fleming et. al. used a conceptually different, though ultimately equivalent, function [28]). If $C$ is a clause and $X$ is a subset of its variables we let $C^X$ denote the subclause of $C$ containing only literals over $X$.

**Definition 6.** Let $C = C_1 \land \cdots \land C_m$ be an unsatisfiable CNF formula and let $(X,Y)$ be any partition of its variables. The unsatisfiability certificate associated with $C$ and $(X,Y)$ is the partial function $\text{cert}^{(X,Y)}_C : \{0,1\}^m \to \{0,1,\ast\}$ defined as

$$\text{cert}^{(X,Y)}_C(\alpha) = \begin{cases} 1 & \text{if } \{C_i^X : C_i \in C, \alpha_i = 0\} \text{ is satisfiable,} \\ 0 & \text{if } \{C_i^Y : C_i \in C, \alpha_i = 1\} \text{ is satisfiable.} \end{cases}$$

When it is clear from context, we may suppress the partition $(X,Y)$ or the underlying CNF formula $C$.

Our second main result is the following monotone feasible interpolation theorem for Sherali-Adams proofs.\(^3\)

\(^3\) A similar feasible interpolation result can be proved for Sum-of-Squares proofs; we follow up on this in an upcoming work.
Theorem 7. Let $C = C_1 \land C_2 \land \cdots \land C_m$ be any unsatisfiable CNF formula. If there is a Sherali-Adams proof of $C$ with size $s$ then for any partition $(X, Y)$ of the variables of $C$ we have $\text{msep}(\text{cert}_C) = O(s^2)$.

Several remarks on this theorem are in order. First, as was shown by Pudlák and Hrubeš [41] a standard interpolation theorem for split formulas follows from this interpolation (simply by “resolving away” the $z$ variables and taking the natural partition of variables). Second, prior to this result there was no monotone feasible interpolation theorem known for Sherali-Adams. A recent work of Hakoniemi gave the first feasible interpolation theorem for Sherali-Adams [37], but, his feasible interpolation result only applied to the “standard” split formulas $A(x, z) \land B(y, z)$ and it only gave a small non-monotone Boolean circuit for the interpolant [37]. Finally, as we have remarked before, the model $\text{msep}(f)$ was recently studied by de Oliveira Olivera and Pudlák under the name weak monotone linear programming gates [20]. They also introduced a model that they called strong monotone linear programming gates, and showed that “circuits” created out of strong monotone linear programming gates can monotone feasibly interpolate Lovász-Schrijver proofs. Pudlák and de Oliveira Olivera left as an open problem whether or not strong monotone linear programming gates can be efficiently simulated by weak monotone linear programming gates. We believe our feasible interpolation result makes this question more interesting, in light of the lack of separations between Lovász-Schrijver and Sherali-Adams proofs – the only separation known is due to Atserias and Ochremiak [3], who showed that degree-6 Lovász-Schrijver has polynomial-size refutations of Tseitin principles (and, furthermore, there are no good size lower bounds on Lovász-Schrijver proofs at all!).

Separation Results

Finally, we prove new separation results between the proof and circuit models described above. In particular, we are interested in the value of $\varepsilon > 0$ that is required in the definition of $\text{msep}(f) := \min_{\varepsilon > 0} \text{rank}^+(S_f^+ - \varepsilon \mathbb{1})$ and in our new characterization of Sherali-Adams degree $\min_{\varepsilon > 0} \text{deg}^+(\text{viol}_C - \varepsilon)$. For instance: is it possible that we never need to take $\varepsilon < 1$? Or, in other words, is it possible that $\text{msep}(f) = \text{msep}_1(f)$, and that Sherali-Adams is already captured by the conical junta degree of $\text{viol}_C - 1$?

We give a negative answer to these questions. First, we show that if $C$ is a pebbling formula, then $\text{deg}^+(\text{viol}_C - 1)$ must be large. (In fact, our lower bound already holds for $\text{deg}^+(\text{viol}_{\text{PEB}_G} - 0.99)$.)

Theorem 8. There is a constant $\delta > 0$ such that for all sufficiently large $m$ there is an in-degree-2 DAG $G$ on $m$ vertices such that

$$\text{deg}^+(\text{viol}_{\text{PEB}_G} - 1) = m^\delta,$$

where $\text{PEB}_G$ is the pebbling formula associated with $G$.

Pebbling formulas are well-known to be refutable in Resolution width $O(1)$ and linear size on graphs of constant in-degree, and thus our result separates $\text{deg}^+(\text{viol}_C - 1)$ from Resolution, resolving an open problem asked by Filmus, Mahajan, Sood and Vinyals [24]. Since Sherali-Adams can efficiently simulate Resolution [18], it follows from combining the previous theorem and Theorem 5 that $\text{deg}^+(\text{viol}_C - 1)$ can be much larger than $\text{deg}^+(\text{viol}_C - \varepsilon)$ for arbitrarily small $\varepsilon > 0$ (although, we note that the standard simulation of Resolution by Sherali-Adams results in $\varepsilon$ that is exponentially small in the size of the proof!).
Next, by using known lifting theorems from conical junta degree to non-negative rank [33,42], we can lift the lower bound in the previous theorem to the following separation between $\text{msep}_1(\text{Gen}_n)$ and $\text{mC}(\text{Gen}_n)$, where $\text{Gen}_n$ is the Generation function.

\textbf{Theorem 9.} The $\text{Gen}_n$ function is computable by polynomial-size monotone circuits, but, there is a $\delta > 0$ such that $\text{msep}_1(\text{Gen}_n) = 2^{\Omega(n^\delta)}$.

Since $\text{msep}(f) \leq \text{mC}(f)$, this provides the analogous separation between $\text{msep}_1(f)$ and $\text{msep}(f)$ (and, indeed, already between $\text{msep}_1(f)$ and monotone circuit size). Taken together, these results imply that choosing smaller $\varepsilon$ actually does increase the power of the corresponding proof systems and circuit models.

1.3 Paper Outline

The rest of this paper is outlined as follows. In Section 2 we carefully outline all important proof systems and non-uniform models of computation that we consider. In Section 3 we prove our new characterization of Sherali-Adams proofs, as well as our feasible interpolation results. In Section 4 we prove our new separations between fragments of Sherali-Adams and between separation complexities.

2 Preliminaries

2.1 Proof Systems

We recall some preliminaries from proof complexity. A clause $C$ is a disjunction of boolean literals; the width of a clause, denoted $w(C)$, is the number of literals in $C$. If $C = C_1 \land C_2 \land \cdots \land C_m$ is a CNF formula then the width of $C$ is the largest width of any clause in $C$.

Resolution Proofs

Fix an unsatisfiable CNF formula $C$ over variables $x_1, \ldots, x_n$. A \textit{Resolution refutation} of $C$ is a sequence of clauses $D_1, \ldots, D_s$ ending in the empty clause $D_s = \emptyset$ such that for each $i \in [s]$, either $D_i$ is in $C$ or is derived from earlier clauses $D_j, D_k$ with $j,k < i$ using the
resolution rule

\[ C \lor \ell, D \lor \overline{\ell} \vdash C \lor D \]

where the rule can only be applied if \( C \) and \( D \) do not contain literals of opposite sign. The size of the resolution proof is \( s \), the number of clauses, and the width of the proof is the largest width of any clause in the proof. Let \( S_{\text{Res}}(C) \) denote the minimum size of any resolution refutation of \( C \), and let \( w_{\text{Res}}(C) \) denote the smallest width of any resolution refutation of \( C \).

Sherali-Adams and Conical Junta Degree

In this paper we will regularly be doing arithmetic with real polynomials that represent boolean functions \( f : \{0, 1\}^n \to \mathbb{R} \). It is well known that each such function can be represented (uniquely) by multilinear polynomials (that is, the largest degree of any individual variable is 1). We will exclusively be operating using multilinear arithmetic with these polynomials – for example, we regard \( x_i^2 \) and \( x_i \) as representing the same boolean function. Formally, we work modulo the ideal \( \langle x_i^2 - x_i \rangle_{i=1}^n \).

A \( d \)-conjunction \( J \) is a conjunction of \( d \) boolean literals. We will usually encode conjunctions as polynomials over the reals, and so if \( S, T \subseteq [n] \) are subsets such that \( S \cap T = \emptyset \) then we let

\[ J_{S,T}(x) := \prod_{i \in S} x_i \prod_{j \in T} (1 - x_j) \]

be the conjunction that tests if all variables in \( S \) are 1 and all variables in \( T \) are 0. Similarly, if \( C \) is a clause, we will let \( \overline{C} \) denote the conjunction such that \( \overline{C}(x) = 1 \) if and only if \( C(x) = 0 \) – that is, \( \overline{C} \) tests if the clause \( C \) is falsified. To spell it out, if \( C = \bigvee_{j \in T} x_j \lor \bigvee_{i \in S} \overline{x}_i \) then \( \overline{C} = J_{S,T} \). Note that the empty conjunction \( J_{\emptyset,\emptyset} = 1 \).

A conical junta is any non-negative real combinations of conjunctions \( J = \sum \lambda_i J_i \), where \( \lambda_i \geq 0 \) is a real constant and \( J_i \) is a conjunction. The degree of a conical junta is the maximum degree of any conjunction in the conical junta representation, and the size of a conical junta is the number of monomials obtained after expanding all juntas but before cancellation. If \( f : \{0, 1\}^n \to \mathbb{R}_{\geq 0} \) is a non-negative real boolean function, then the conical junta degree of \( f \), denoted \( \deg^+(f) \), is the minimum integer \( d \) such that we can write

\[ f = \sum \lambda_i J_i \]

where \( \lambda_i \geq 0 \) is a non-negative real constant and \( J_i \) is a \( \leq d \)-conjunction for each \( i \). Note that, unlike polynomial degree, the representation of \( f \) as a conical junta is not necessarily unique.

A Sherali-Adams refutation of an unsatisfiable CNF formula \( C = C_1 \land \cdots \land C_m \) is given by a sequence of conical juntas \( J_1, J_2, \ldots, J_m, J_{m+1} \) such that

\[ \sum_{i=1}^m -\overline{C}_i J_i + J_{m+1} = -1. \]

The degree of the proof is the maximum degree of any of the polynomials \( -\overline{C}_i J_i, J \) before cancellation. The size of the proof is the number of distinct monomials that occurs in the proof (after expanding all polynomials and before cancellation). Let \( \text{SA}(C) \) denote the minimum degree of any Sherali-Adams refutation of \( C \). We note that CNF formulas are often encoded as linear inequalities rather than juntas; this is equivalent to the above definition.
up to increasing the degree by an additive factor of $w(C)$ (see e.g., [27, Claim 3.32]). If $C = C_1 \land C_2 \land \cdots \land C_m$ is a CNF formula then define $\text{viol}_C : \{0,1\}^n \to \mathbb{R}_{\geq 0}$ by

$$\text{viol}_C(x) := \sum_{i=1}^{m} \overline{C}_i(x).$$

That is, $\text{viol}_C(x)$ counts the number of falsified clauses of $C$ on input $x$. It follows that $C$ is unsatisfiable if and only if $\text{viol}_C(x) \geq 1$ for all $x \in \{0,1\}^n$ – indeed, if and only if $\text{viol}_C(x) \geq \varepsilon$ for any $\varepsilon > 0$. Thus, a conical junta representation of $\text{viol}_C - \varepsilon$ for any $\varepsilon > 0$ constitutes a proof that $C$ is unsatisfiable.

In fact, conical junta representations of $\text{viol}_C - \varepsilon$ correspond naturally to fragments of Sherali-Adams proofs. This is easy to see: suppose that $\text{viol}_C - \varepsilon = J$ where $J$ is a conical junta. Explicitly, this means that

$$\sum_{j=1}^{m} \frac{1}{\varepsilon} \overline{C}_j + \frac{J}{\varepsilon} = -1.$$

Conversely, suppose that we have a Sherali-Adams refutation of $C$ of the form

$$\sum_{j=1}^{m} -\alpha_j \overline{C}_j + J = -1$$

where $0 \leq \alpha_j \leq 1/\varepsilon$ for all $j$. Since $\overline{C}_j$ is a conjunction and $1/\varepsilon \geq \alpha_j$, this means that

$$\sum_{j=1}^{m} \frac{1}{\varepsilon} \overline{C}_j + J + \sum_{j=1}^{m} \left(\frac{1}{\varepsilon} - \alpha_j\right) \overline{C}_j = -1$$

is also a Sherali-Adams refutation of $C$. Multiplying through by $\varepsilon$ yields the following:

\begin{lemma}
Let $C = C_1 \land C_2 \land \cdots \land C_m$ be an unsatisfiable CNF formula. If there is a Sherali-Adams refutation of the form

$$\sum_{j=1}^{m} -\alpha_j \overline{C}_j + J = -1$$

where $J$ is a conical junta of degree $d$ and size $s$ and with $0 \leq \alpha_j \leq 1/\varepsilon$ for each $\alpha_j$, then there is a degree-$\max\{d, w(C)\}$ conical junta $J'$ such that $\text{viol}_C - \varepsilon = J'$.

Conversely, if $\text{viol}_C - \varepsilon = J'$ for some degree-$d$, size $s$ conical junta $J'$, then there is a Sherali-Adams proof of the above form with $\alpha_j \leq 1/\varepsilon$ and where $J = J'$.
\end{lemma}

As we will show in Subsection 3.1, it turns out that this fragment of Sherali-Adams is actually complete, in the sense that any Sherali-Adams proof can be put into this form without changing the degree and without (badly) changing the size.
2.2 Circuit Complexity and Extended Formulations

Boolean Circuit Complexity

A partial boolean function is a function \( f : \{0, 1\}^n \to \{0, 1, *\} \) (the * represents that we “don’t care” what the function’s output is). Given two partial boolean functions \( f, g \) we say that \( g \) extends \( f \) if for all \( x \) such that \( f(x) \neq * \), \( g(x) = f(x) \). A (total) boolean function \( f : \{0, 1\}^n \to \{0, 1\} \) is monotone if \( x \leq y \Rightarrow f(x) \leq f(y) \) (where the first inequality is taken coordinate-wise); a partial boolean function \( f : \{0, 1\}^n \to \{0, 1, *\} \) is monotone if there is a total monotone boolean function \( g \) extending \( f \).

We assume familiarity with standard boolean circuit models in complexity theory – e.g. boolean formulas, boolean circuits, etc. All circuits considered in this paper will be composed of AND (\( \land \)) and OR (\( \lor \)) gates, and will be De-Morgan in the sense that if any negations appear they appear at the inputs of the circuit. A boolean circuit is monotone if it does not contain any negation gates. We say that a circuit \( C \) computes a partial function \( f : \{0, 1\}^n \to \{0, 1, *\} \) if \( f(x) = 1 \Rightarrow C(x) = 1 \) and \( f(x) = 0 \Rightarrow C(x) = 0 \) for all \( x \in \{0, 1\}^n \) (in other words, the total function computed by \( C \) extends \( f \)). For any partial boolean \( f \) we let

- \( \text{C}(f) := \) the size of the smallest boolean circuit computing \( f \),
- \( \text{mC}(f) := \) the size of the smallest monotone boolean circuit computing \( f \),
- \( F(f) := \) the size of the smallest boolean formula computing \( f \),
- \( \text{mF}(f) := \) the size of the smallest monotone boolean formula computing \( f \).

Non-negative Rank and Extended Formulations

Let \( B \) be a non-negative real matrix. The non-negative rank of \( B \), denoted rank\(^+\)(\( B \)), is the smallest integer \( r \) such that \( B \) can be written as the sum of \( r \) rank-1 non-negative matrices.

Non-negative rank is a very well-studied parameter in theoretical computer science, owing to its close relationship with polytopes in convex optimization. We, however, will be particularly interested in the connection between non-negative rank and circuit complexity.

Let \( I_{m,n} \) denote the \( m \times n \) all-1s matrix (we may suppress the \( m, n \) when it is clear from context). Let \( f : \{0, 1\}^n \to \{0, 1, *\} \) be a partial boolean function, let \( U = f^{-1}(1) \) and let \( V = f^{-1}(0) \). We let \( S_f \) be the \( |U| \times |V| \) matrix defined by, for each \( x \in U, y \in V \),

\[
S_f(x, y) = \sum_{i=1}^{n} [x_i \neq y_i]
\]

where \([P]\) is the \( \{0, 1\}\)-indicator function for the predicate \( P \). That is, the matrix \( S_f \) counts the number of witnesses that \( x \neq y \) for each \( x \in f^{-1}(1), y \in f^{-1}(0) \). If the function \( f \) is furthermore monotone, we define the matrix \( S_f^+ \) by

\[
S_f^+ = \sum_{i=1}^{n} [x_i = 1 \land y_i = 0].
\]

Observe that both \( S_f(x, y), S_f^+(x, y) \geq 1 \) for all \( x \in U, y \in V \). Because of this, we can say that for any \( 0 < \varepsilon \leq 1 \) the matrices \( S_f - \varepsilon I \) and \( S_f^+ - \varepsilon I \) are both non-negative. With this in mind, we make the following definitions (the second two, of course, only make sense for
monotone $f$):
\[
\begin{align*}
\text{sep}_1(f) & := \text{rank}^+(S_f - \mathds{1}) \\
\text{sep}(f) & := \min_{\varepsilon > 0} \text{rank}^+(S_f - \varepsilon \mathds{1}) \\
\text{msep}_1(f) & := \text{rank}^+(S_f^+ - \mathds{1}) \\
\text{msep}(f) & := \min_{\varepsilon > 0} \text{rank}^+(S_f^+ - \varepsilon \mathds{1})
\end{align*}
\]

We note the similarity between the definition of $\text{sep}(f)$ and $\deg^+(\text{viol}_C - \varepsilon)$ – the restriction of Sherali-Adams – which was introduced in the previous section. Furthermore, as we have discussed in detail in Section 1, each of these parameters can be interpreted as the minimum size of any extended formulation separating 0s of a function from 1s of a function, and they are closely related to circuit and formula size (cf. Theorem 3, Theorem 4).

3 Upper Bounds

3.1 A New Characterization of Sherali-Adams Degree

In this section we provide our new characterization of Sherali-Adams degree, which is recorded next.

\textbf{Theorem 5.} For any unsatisfiable CNF formula $C$, if there is a Sherali-Adams refutation of $C$ with degree $d$ and size $s$, then there is an $\varepsilon > 0$ and a degree $d$, size $s$ conical junta $J$ such that
\[
\text{viol}_C - \varepsilon = J.
\]

Consequently, if $SA(C)$ is the minimum degree of any Sherali-Adams refutation of $C$, then
\[
SA(C) = \min_{\varepsilon > 0} \deg^+(\text{viol}_C - \varepsilon).
\]

Proof. Let
\[
\sum_{i=1}^{m} -\overline{C}_i J_i + \sum_j \lambda_j J_j = -1
\]
be a degree-$d$, size-$s$ Sherali-Adams refutation of $C$, where each $J_i$ are conical juntas. By expanding each of the $J_i$ into their monomials, we can re-write this proof as
\[
\sum_{k=1}^{t} -\alpha_k \overline{C}_{i_k} m_k + \sum_j \lambda_j J_j = -1 \tag{1}
\]
for some $t$ where $\alpha_k \geq 0$ is a non-negative real constant and $m_k$ is a monomial for each $k$.

Now, fix some $1 \leq u \leq t$. Assume w.l.o.g. that the variables occurring in $m_u$ are $x_1, x_2, \ldots, x_\ell$ for some $\ell$ – explicitly, since $m_u$ is a monomial we therefore have $m_u = \prod_{i=1}^{\ell} x_i = J_{[\ell],\emptyset}$. The main claim we use is the following.

\textbf{Claim.} Let $x_1, x_2, \ldots, x_n$ be any set of boolean variables. Then
\[
\sum_{S \subseteq [n]} J_{S, [n] \setminus S} = 1.
\]
Proof of Claim. By induction on \( n \). If \( n = 0 \) then the only term in the sum is \( J_{\emptyset, \emptyset} \), which is 1. If \( n > 0 \), then
\[
\sum_{S \subseteq \{\ell\}, \ell \neq \ell'} J_{S, \{\ell\} \setminus S} + \sum_{S \subseteq \{\ell\}} J_{S, \{\ell\} \setminus S} = x_n (\sum_{S \subseteq \{\ell\}, \ell \neq \ell'} J_{S, \{\ell\} \setminus S} (1 - x_n)) + (1 - x_n) (\sum_{S \subseteq \{\ell\}} J_{S, \{\ell\} \setminus S}) = x_1 + (1 - x_1) = 1 \]
where we applied the induction hypothesis twice. \( \square \)

Define the conical junta
\[
J_u := \sum_{S \subseteq \{\ell\}, \ell \neq \ell'} J_{S, \{\ell\} \setminus S} \equiv \sum_{S \subseteq \{\ell\}} J_{S, \{\ell\} \setminus S} - J_{\emptyset, \emptyset}.
\]

By the Claim and the definition of \( J_u \) we have \( m_u + J_u = J_{\emptyset, \emptyset} + J_u = 1 \). This means that
\[
-\alpha_u m_u \overline{c_{i_u}} = -\alpha_u \overline{c_{i_u}} J_{\emptyset, \emptyset} + \alpha_u \overline{c_{i_u}} J_u = \alpha_u \overline{c_{i_u}} J_u
\]

Since \( \alpha_u > 0 \) and \( \overline{c_{i_u}} \) is a conjunction it follows that \( \alpha_u \overline{c_{i_u}} J_u \) is a conical junta. By repeating this construction for every polynomial \(-\alpha_k \overline{C_{i_k}} m_k \) it follows that we can rewrite the Sherali-Adams proof from Equation 1 as
\[
\sum_{k=1}^t -\alpha_k \overline{C_{i_k}} m_k + \left( \sum_j \lambda_j J_j + \sum_{k=1}^t \alpha_k \overline{C_{i_k}} J_k \right) = -1.
\]

Replacing \(-\alpha_k \overline{C_{i_k}} m_k \) by \( \alpha_k C_{i_k} J_k \) preserves the degree but increases the monomial size of the proof by a \( 2^d \) factor. By applying Lemma 10 it follows that there is a \( \varepsilon > 0 \) such that \( \text{viol}_C - \varepsilon = J \) for some size at most \( s 2^d \), degree-\( d \) conical junta \( J \). \( \square \)

### 3.2 Feasible Interpolation for Sherali-Adams

In this section, we give a monotone interpolation theorem for Sherali-Adams in terms of the unsatisfiability certificate (cf. Definition 6). For our purposes, the important property of the unsatisfiability certificate is the following connection it shares with \( \text{viol}_C \).

**Lemma 11.** Let \( C \) be an unsatisfiable CNF formula on \( m \) clauses, let \( (X, Y) \) be any partition of the variables of \( C \), and let \( \text{cert} \) be the corresponding unsatisfiability certificate. There are surjections \( \mu : \{0, 1\}^X \to \text{cert}^{-1}(1) \) and \( \nu : \{0, 1\}^Y \to \text{cert}^{-1}(0) \) such that
\[
S_{\text{cert}}^+(\mu(x), \nu(y)) = \text{viol}_C(xy)
\]
where \( xy \) denotes the concatenation of \( x, y \).

**Proof.** We define \( \mu \) and \( \nu \) by their evaluations on assignments to the \( X \) and \( Y \) variables as follows. For any \( x \in \{0, 1\}^X \) define \( \mu^x \in \{0, 1\}^m \) by \( \mu^x_i = 1 - C_i^X(x) \) — that is, \( \mu^x_i \) is 1 iff \( C_i^X(x) \) is falsified. Similarly, for any \( y \in \{0, 1\}^Y \) define \( \nu^y \in \{0, 1\}^m \) by \( \nu^y_i = C_i^Y(y) \) — so, \( \nu^y_i \) is 1 iff \( C_i^Y(y) \) is satisfied. It follows by definition that \( \text{cert}(\mu^x) = 1 \) and \( \text{cert}(\nu^y) = 0 \), and furthermore it is clear that \( x \mapsto \mu^x \) and \( y \mapsto \nu^y \) are surjective maps since every \( 1 \) and \( 0 \) assignment to \( \text{cert} \) must have some witnessing assignments to the underlying variables. Finally, observe that if \( \mu^x_i = 1 \) and \( \nu^y_i = 0 \) then \( C_i^X(x) = C_i^Y(y) = 0 \), and thus the combined assignment \( xy \) violates the clause \( C_i \). The converse also clearly holds, proving the lemma. \( \square \)

By combining this fact with Theorem 5, we can immediately prove the following monotone feasible interpolation theorem for Sherali-Adams proofs. This next theorem is slightly weaker (in that we lose a \( 2^d \) factor in the size) than the interpolation theorem stated in Section 1, but, has the benefit of admitting a completely transparent proof. We show how to remove the \( 2^d \) factor (thus proving Theorem 7) in the full version of this paper; we do not include it here due to space constraints.
Theorem 12. For any unsatisfiable CNF $C$ and any partition of its variables $(X, Y)$, if there is a conical junta $J$ of size $s$ and an $\varepsilon > 0$ such that $\text{viol}_C - \varepsilon = J$ then
\[
\text{rank}^+(\text{cert}_C - \varepsilon) \leq s.
\]
Consequently, if there is a degree-$d$, size-$s$ Sherali-Adams refutation of $C$ then $\text{msep}(\text{cert}_C) \leq s2^d$.

Proof. The “Consequently” statement of the theorem is just an application of Theorem 5, so we focus on proving the non-negative rank upper bound for $\text{cert}_C$.

Write $\text{viol}_C - \varepsilon = \sum \lambda_i J_i$ where each $J_i$ is a conjunction. Under the partition of variables $(X, Y)$, we regard this as a $\{0, 1\}^X \times \{0, 1\}^Y$ matrix, where the $(x, y)$th entry of the matrix is exactly $\text{viol}_C(xy) - \varepsilon$. By Lemma 11 it follows immediately that $\text{rank}^+(\text{cert}_C - \varepsilon) = \text{rank}^+(\text{vio}_C - \varepsilon)$, since the existence of the two surjections implies that the matrix $\text{viol}_C - \varepsilon$ is exactly the matrix $\text{cert}_C - \varepsilon$ with some extra copies of rows and columns padded in. Thus if we show $\text{rank}^+(\text{viol}_C - \varepsilon) \leq s$ the proof of the theorem will be completed.

To show this, we use the expression of $\text{viol}_C - \varepsilon$ as a conical junta. Observe that if $J_i$ is a conjunction over $X$ and $Y$ variables, then we can interpret any term $z_i$ or $1 - z_i$ in the product $J_i$ as a rank-1 $2^{|X|} \times 2^{|Y|}$ non-negative matrix (indeed, it is rank-1 since each term depends only on the $x$ assignment or the $y$ assignment). If we let $A \circ B$ denote the Hadamard (i.e. entrywise) product of two matrices $A$ and $B$, it follows that the conjunction $J_i$ is simply the Hadamard product of a number of rank-1 non-negative matrices. Since $\text{rank}^+(A \circ B) \leq \text{rank}^+(A) \text{rank}^+(B)$ – a fact easily verified by taking non-negative rank decompositions – it follows that $\text{rank}^+(J_i) = 1$. Therefore, if we can write $\text{viol}_C - \varepsilon = J$ where $J$ is a non-negative combination of $s$ conjunctions, it immediately follows by the subadditivity of rank$^+$ that $\text{rank}^+(\text{viol}_C - \varepsilon) \leq s$, proving the theorem.

4 Lower Bounds

In this section we prove our new separation between fragments of Sherali-Adams and Resolution and then show how to lift this separation to a new separation between extension complexity (particularly, $\text{msep}_1$) and monotone circuit complexity. First, we show that the well-studied Pebbling formulas $\text{Peb}_G$ require large conical junta degree to represent $\text{viol}_{\text{Peb}_G} - 1$ (cf. Theorem 8). However, our eventual goal is to apply a lifting theorem to get a strong lower bound on $\text{msep}_1(f) := \text{rank}^+(S_f^+ - 1)$ from the lower bound on $\text{deg}^+_{\varepsilon}(\text{viol}_{\text{Peb}_G} - 1)$ for some specially tailored monotone boolean function $f$. Unfortunately, the known lifting theorems from conical junta degree to non-negative rank [33,42] only apply to approximate conical junta degree, which complicates the lower bound slightly.

Definition 13 (Approximate conical junta degree). For a non-negative function $f : \{0, 1\}^n \rightarrow \mathbb{R}_{\geq 0}$ we define its $\varepsilon$-approximate conical junta degree $\text{deg}^+_{\varepsilon}(f)$, as the minimum degree of a conical junta $J$ such that $J(x) \in f(x) \pm \varepsilon$ for all inputs $x$.

For convenience, instead of proving lower bounds on the conical junta degree of $\text{Peb}_G$ directly, we will instead work with the slightly more flexible Iteration problem $\text{Iter}_n$ (this is also sometimes known as the Housesitting Principle) [4,12]. An easy reduction from $\text{Iter}_n$ to $\text{Peb}_G$ will then yield the degree lower bound for Pebbling formulas.

4.1 Conical Junta Lower Bound for Iteration

We study the Iteration problem $\text{Iter}_n$ on an $n$-by-$n$ grid. An input to this problem consists of $n^2$ input variables $x_v$, one for each grid node $v \in [n] \times [n]$. The variables take values from an alphabet of size $O(n)$. Namely, for each $v \in [n-1] \times [n]$, we have $x_v \in [n] \cup \{\perp\}$. If
$x_v \in [n]$, we say $v$ is \textit{active}, and it points to the $x_v$-th node on the next row. If $x_v = \perp$, we say $v$ is \textit{disabled}. For each node on the final row $v \in \{\top, \perp\}$ there is a boolean variable $x_v \in \{\top, \perp\}$ where $x_v = \top$ means $x_v$ is \textit{active} and $x_v = \perp$ means $v$ is \textit{disabled}. The goal is to output one of the following types of solution.

(i) the \textit{distinguished source node} $(1,1)$ is disabled ($x_{1,1} = \perp$), or

(ii) an active node that points to a disabled node ($x_{i,j} = k \in [n]$ and $x_{i+1,k} = \perp$), or

(iii) an active node on the final row ($x_{n,j} = \top$).

For larger-than-boolean alphabets, we generalise the definition of a conjunction $J$ naturally to be a partial assignment of pointers to nodes; for example, $J = [x_{1,1} = 3; x_{2,3} = \perp; x_{4,1} = \top]$ is a conjunction of degree 3. We prove strong lower bounds against the conical junta degree of $\text{viol}_{\text{Iter}_n} - 1$; indeed, we are even able to prove lower bounds against the conical junta degree of $\text{viol}_{\text{Iter}_n} - 0.99$, which will turn out to be crucial for later results.

\textbf{Theorem 14.} $\deg^+_\text{Iter}_n(\text{viol}_{\text{Iter}_n} - 1) \geq \Omega(n)$.

\textbf{Proof.} We define three random inputs $X_1, X_2, X_3$ that admit 1, 2, or 3 solutions, all of type (iii). The inputs are constructed from given data $(v_i')$, $j \in [2]$, $i \in [n]$, that specifies for each row $i \in [n]$, two (or three in the case of $X_3$) nodes $v_i^1, v_i^2 \in \{\top, \perp\}$ at random subject to $v_i^1 = (1,1), v_i^1 \neq v_i^2$ for $i \in [n-1], v_n^1 = v_n^2$. Then $X_1 = X(v_i')$.

$X_2$ = \textquotedblleft two random disjoint paths.	extquotedblright  We choose $v_i^j$ for $j \in [2], i \in [n]$ uniformly at random subject to $v_i^1 = (1,1), v_i^1 \neq v_i^2$ for $i \in [n]$. Then $X_2 = X(v_i')$.

$X_3$ = \textquotedblleft three random disjoint paths.	extquotedblright  We choose $v_i^j$ for $j \in [3], i \in [n]$ uniformly at random subject to $v_i^1 = (1,1), v_i^1 \neq v_i^j$ for $j \neq j', i \in [n]$. Then $X_3 = X(v_i')$.

Suppose for the sake of contradiction that $J = \sum_i \lambda_i J_i$ is a degree-$o(n)$ conical junta that for every input $x$ outputs $J(x) \in \text{viol}_{\text{Iter}_n}(x) - 1 \pm \epsilon$ where $\epsilon := 1/20$. We may assume that each conjunction $J$ in $J$ is such that if it reads an active node $v$ on row $n - 1$ (meaning that the conjunction contains $[x_{n-1,j} = k]$ for some $k$ and $v$ is node $(n-1,j)$) then it also reads the boolean variable associated with node $(n, x_n)$ on the last row. (Indeed, we may always replace $J$ containing $[x_{n-1,j} = k]$ with the equivalent conic combination $J \cdot [x_{n,k} = \top] + J \cdot [x_{n,k} = \perp]$.)
We call a conjunction \( J \) paired if it reads at least two active nodes on the last row, that is, \( J \) witnesses at least two solutions of type (iii).

▷ Claim 15. If \( J \) is not paired, then \( \mathbb{E}[J(X_2)] \leq 2 \cdot \mathbb{E}[J(X_1)] \).

▷ Claim 16. If \( J \) is paired, then \( \mathbb{E}[J(X_3)] \geq (3 - o(1)) \cdot \mathbb{E}[J(X_2)] \).

We prove the claims shortly after we first complete the proof of the theorem assuming them. We write \( J = \sum J_i = J_{\text{not}} + J_{\text{pair}} \) where \( J_{\text{pair}} \) consists of conjunctions that are paired and \( J_{\text{not}} \) consists of those that are not. We calculate

\[
1 - \epsilon \leq \mathbb{E}[J(X_2)] = \mathbb{E}[J_{\text{not}}(X_2)] + \mathbb{E}[J_{\text{pair}}(X_2)] \\
\leq 2\mathbb{E}[J_{\text{not}}(X_1)] + \mathbb{E}[J_{\text{pair}}(X_2)] \quad \text{(Claim 15 and linearity of } \mathbb{E}) \\
\leq 2\epsilon + \mathbb{E}[J_{\text{pair}}(X_3)]/(3 - o(1)) \quad \text{(Claim 16 and linearity of } \mathbb{E}) \\
\leq 2\epsilon + (2 + \epsilon)/(3 - o(1)) \quad \text{ } \quad \text{(} \mathcal{J} \text{ approximates } \text{viol} - 1) \\
< 1 - \epsilon.
\]

This is the desired contradiction that concludes the proof (modulo the Claims).

Proof of Claim 15. The claim says that \( \Pr[J(X_1) = 1] \geq \Pr[J(X_2) = 1]/2 \) for a non-paired \( J \). Consider any fixed setting of the \( v_t^r \) such that \( X(v_t^r) \) is in the support of \( X_2 \) and such that \( J \) accepts \( X(v_t^r) \). Since \( J \) is not paired, it does not read the nodes \( \{v_t^{b, i} \}_{i \in [n]} \) for some \( b \in [2] \), and hence if we pick \( b \in [2] \) uniformly at random, \( J_{\text{pair}}[J((v_t^{b, i}))_{i \in [n]}, v_t^{b, i}, v_t^{b, i} = 1] \geq 1/2 \). Moreover, if \( (v_t^{b, i})_{i \in [n]} \) is distributed as in the process for \( X_2 \), then \( ((v_t^{b, i}))_{i \in [n]}, v_t^{b, i}, v_t^{b, i} \) (for random \( b \in [2] \)) is distributed as in the process for \( X_1 \). Hence \( J \)'s acceptance probability can decrease by at most a factor of \( 1/2 \).

Proof of Claim 16. The claim says that \( \Pr[J(X_3) = 1] \geq (3 - o(1)) \Pr[J(X_2) = 1] \) for a paired \( J \). Since \( J \) reads at most \( o(n) \) nodes, there is some middle row \( i^* \in [n/3, 2n/3] \) from which \( J \) reads no nodes. Split any input \( x \) into three parts so that \( x = (x^r, x^i, x^b) \) where \( x^r \) is the pointer assignment for nodes in topmost \( i^* \) rows, \( x^i \) is for nodes in row \( i^* \), and \( x^b \) is for nodes in the remaining bottommost rows. Write also \( J(x) = J^r(x^r) \cdot J^i(x^i) \cdot J^b(x^b) \). Note that for both \( i = 2, 3 \), the variables \( X_i^r \) and \( X_i^b \) are independent and thus

\[
\Pr[J(X_i) = 1] = \Pr[J^r(X_i^r) = 1 \land J^i(X_i^i) = 1] = \Pr[J^r(X_i^r) = 1] \cdot \Pr[J^b(X_i^b) = 1]. \tag{2}
\]

We will prove the following estimates, which, when plugged into (2), would complete the proof:

\[
\Pr[J^r(X_3^r) = 1] \geq (1 - o(1)) \Pr[J^r(X_2^r) = 1], \tag{3}
\]

\[
\Pr[J^b(X_3^b) = 1] \geq (3 - o(1)) \Pr[J^b(X_2^b) = 1]. \tag{4}
\]

Let us prove (3). Consider generating a sample from \( X_3^r \) as follows: (i) sample \( x^r \sim X_2^r \), and (ii) add to \( x^r \) a random third path which is disjoint from the existing two. We argue that if the sample \( x^r \) in step (i) is accepted by \( J^r \), then the conjunction continues to accept after step (ii) with high probability. Indeed, conditioned on the first step being accepting for \( J^r \), we note that the probability that any one node is picked to lie on the third path is at most \( 1/(n - 2) \) and hence (by a union bound) the probability that some node read by \( J^r \) lies
on the third path is at most \( \text{deg}(J)/(n-2) = o(n)/(n-2) = o(1) \). Hence \( \tilde{J} \) continues to accept with probability \( 1 - o(1) \), which proves (3).

Let us finally prove (4). Note how \( X^b_3 \) (resp. \( X^b_3 \)) is distributed: it consists of two (three) uniformly random disjoint paths down the grid. Consider the following bipartite graph \( (L \cup R, E) \) where

- Left vertices \( L \) are outcomes of \( X^b_3 \) (pairs of disjoint paths).
- Right vertices \( R \) are outcomes of \( X^o_3 \) (triples of disjoint paths).
- There is an edge \((l, r) \in E\) iff the two paths of \( l \) are a subset of the three paths of \( r \).

Note that this graph is biregular with right degree \( d_R := 3|R|/|L| \).

Denote by \( L' \subseteq L \) (resp. \( R' \subseteq R \)) the set of \( l \in L \) (resp. \( r \in R \)) accepted by \( J^b \). Let \( E(L', R') := L' \times R' \cap E \) denote the set of edges between \( L' \) and \( R' \). Note that if \( r \in R' \), then at most one of \( r \)'s neighbours is in \( L' \); this is because \( J^b \) is paired and hence it requires a prescribed pair of the three paths in \( r \) to be present. Consequently \( |E(L', R')| \leq |R'| \).

On the other hand, if \( l \in L' \), then \( 1 - o(1) \) fraction of its neighbours are in \( R' \); this is because of the same argument as in the preceding paragraph (condition on \( X^o_3 = l \in L' \) in step (i) and choose a random neighbour of \( l \) in step (ii)). Hence \( |E(L', R')| \geq (1 - o(1))d_L|L'| \). We now put these observations together to prove (4):

\[
\Pr[J^b(X^b_3) = 1] = \frac{|R'|/|R|}{|E(L', R')|/|R|} \\
\geq \frac{(1 - o(1))d_L|L'|/|R|}{(3 - o(1))|L'|/|L|} \\
= \frac{(3 - o(1))\Pr[J^o(X^o_3)]}{1} = 1.
\]

4.2 Lower bound for Pebbling

The **Pebbling problem** \( \text{Peb}_G \) is defined relative to a DAG \( G = (V, E, v^*) \) where \( v^* \) is a **distinguished source node** (in-degree 0) and where every node has out-degree at most 2 (but in-degree may be unbounded). The input to \( \text{Peb}_G \) is \( y \in \{0,1\}^E \), that is, an boolean assignment to the edges of \( G \). Such an assignment \( y \) naturally defines a subgraph \( G_y \) of \( G \) consisting of all the edges \( e \) such that \( y_e = 1 \), which we call the **active** edges. The goal is to output one of the following types of solution.

(i) the distinguished source node is a sink (out-degree 0) in \( G_y \).

(ii) a node \( v \in V \) that is a **proper sink** in \( G_y \), meaning \( v \) has in-degree \( \geq 1 \) and out-degree 0.

Note that the presence of any solution can be certified by reading at most 3 bits, and hence \( \text{Peb}_G \) corresponds to the canonical search problem of an unsatisfiable 3-CNF formula. Moreover, if \( G \) has \( m \) edges then this 3-CNF formula has at most \( m + 1 \) clauses, as every possible violation in \( \text{Peb}_G \) is either the violation of the first type, or can be identified with an incoming edge to a node. It follows that for all \( z \in \{0,1\}^m \)

\[
\text{viol}_{\text{Peb}_G}(z) - 1 \leq m,
\]

which we have recorded for later use.

We prove the following theorem, which is a strengthening of Theorem 8.

**Theorem 17.** There is a DAG \( G \) with \( m \) edges such that \( \deg_{1/20}^{-1}(\text{viol}_{\text{Peb}_G} - 1) \geq m^{\Omega(1)} \).
As we have seen in Section 2, if $f$:

\[
mC(f) \leq \text{msep}(f) \leq \text{msep}_1(f) \leq \text{mF}(f),
\]

where $mC(f)$ (mF(f)) is the size of the smallest monotone circuit (formula, resp.) computing $f$. It is natural to wonder if it is necessary to choose $\varepsilon < 1$ in order to simulate monotone circuits – in other words, whether or not msep$_1(f)$ is already upper bounded by mC(f). In this section, we show that the separation proven in Subsection 4.2 implies that such a bound is not possible. In particular, we prove the following theorem:

\begin{itemize}
  \item \textbf{Theorem 9.} The Gen$_n$ function is computable by polynomial-size monotone circuits, but, there is a $\delta > 0$ such that msep$_1(Gen_n) = 2^{\Omega(n^\delta)}$.
\end{itemize}

Let us begin by defining the Gen$_n$ function, which was originally introduced by Raz and McKenzie [50] and has now been the subject of a number of works in circuit complexity [13, 22, 23, 34, 50, 53].

\begin{itemize}
  \item \textbf{Definition 18.} Let $n$ be a positive integer. A set of triples $T \subseteq [n]^3$ is said to generate $k \in [n]$ if $k = 1$ or if there is a triple $(i, j, k) \in T$ such that $T$ generates both $i$ and $j$.
\end{itemize}
The $\text{GEN}_n$ function is a monotone boolean function defined as follows. As input, the function receives a list of triples $T \subseteq \{0,1\}^3$ encoded as a binary string of length $\{0,1\}^n$. It then outputs 1 if and only if $n$ can be generated from $T$.

It is well known that $\text{GEN}_n$ has polynomial-size monotone circuits [50], so, we focus on proving the lower bound in Theorem 9. To do this we apply the following lifting theorem from conical junta degree to non-negative rank due to Kothari, Meka, and Raghavendra [42] (we note that applying the lifting theorem from Göös et. al. would also yield an essentially identical result [33]). In order to state the theorem we introduce the important notion of a pattern matrix:

**Definition 19.** Let $f : \{0,1\}^n \rightarrow \mathbb{R}$ be a boolean function, and let $g : \mathcal{X} \times \mathcal{Y} \rightarrow \{0,1\}$ be any function. Define the pattern matrix $f \circ g^m : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$ for any $x \in \mathcal{X}^n, y \in \mathcal{Y}^n$ by

$$(f \circ g^m)(x, y) := f(g(x_1, y_1), g(x_2, y_2), \ldots, g(x_n, y_n)).$$

We note that even though we have defined it as a function, we think of $f \circ g^m$ as a $\mathcal{X}^n \times \mathcal{Y}^n$ matrix in the natural way.

The next theorem is a slight modification of Theorem 1.7 in [42], and follows immediately from the proof of [42].

**Theorem 20.** Let $f : \{0,1\}^n \rightarrow \mathbb{R}_{\geq 0}$ be any boolean function with $\mathbb{E}[f] = 1$, and suppose that $\deg_+^+(f + 1/100n) \geq 9 \deg(f)$. For any $b \geq 100 \log n$ there is a gadget function $g : \{0,1\}^b \times \{0,1\}^b \rightarrow \{0,1\}$ such that

$$\text{rank}_+^+(f \circ g^m) = 2^{\Omega(b \deg_+^+(f + 1/100n))}.$$ 

We combine this theorem with our lower bound for $\text{PEB}_G$ from the previous section (Theorem 17) and an embedding argument to prove the Theorem 9.

**Theorem 21.** There is a constant $\delta > 0$ such that for all sufficiently large $n$, $\text{msep}_1(\text{GEN}_n) \geq 2^{\Omega(n^\delta)}$.

**Proof.** Let $G$ be the DAG with $m$ edges guaranteeing the approximate conical junta degree lower bound from Theorem 17, and let $g$ be the gadget function guaranteed by Theorem 20. We begin by showing that

$$\text{rank}_+^+((\text{viol}_{\text{PEB}_G} - 1) \circ g^m) = 2^{\Omega(m^\delta)}$$

for some universal constant $\delta > 0$. Once we have this, we show that the pattern matrix $(\text{viol}_{\text{PEB}_G} - 1) \circ g^m$ can be embedded as a submatrix of the slack matrix $S^T_{U,V}$ associated with $\text{GEN}_n$ (cf. Section 2); since non-negative rank is monotone decreasing with respect to taking submatrices this immediately implies $\text{msep}_1(\text{GEN}_n) \geq 2^{\Omega(n^\delta)}$, proving the theorem.

Let $f = \text{viol}_{\text{PEB}_G} - 1$, with the goal of applying the previous lifting theorem, and note $f : \{0,1\}^m \rightarrow \mathbb{R}_{\geq 0}$ and furthermore that $f(x) \leq m$ by Equation 5. It follows that $\mathbb{E}f \leq m$, and thus there is some $\varepsilon \leq 1/100$ such that

$$\deg_+^+(f/\mathbb{E}f + 1/100m) = \deg_+^+(\text{viol}_{\text{PEB}_G} - 1 + \mathbb{E}f/100m)$$

$$= \deg_+^+(\text{viol}_{\text{PEB}_G} - (1 - \varepsilon))$$

$$\geq \deg_+^+(\text{viol}_{\text{PEB}_G} - 1) \geq m^\delta$$

where we have used the fact that $\deg_+^+$ is invariant under multiplying by positive constants, and also Theorem 17. Finally, since $\text{viol}_{\text{PEB}_G}$ can be represented as a sum of violations of a
We define an $n$-CNF formula it follows that $\deg(f) = 3$. We can therefore apply Theorem 20 and conclude that

$$\text{rank}^+((\text{viol}_{\text{PEG}} - 1) \circ g^m) = 2^\Omega(m^k).$$

So, in the remainder of the proof we show, for any DAG $G$ with $m$ edges, $(\text{viol}_{\text{PEG}} - 1) \circ g^m$ can be embedded as a submatrix of the slack matrix $S^+_\text{PEG}$ associated with the $\text{PEG}_n$ function for some $n = \text{poly}(m)$. This is the same embedding argument that has been used in many works relating “lifted” PEBG to $\text{Gen}_n$ [22,34,50,53]; for this reason, we only sketch the argument for the sake of completeness. Recall from Section 2 the definition of $S^+_f$: if $f : \{0,1,*\}^n \to \{0,1\}$ is a partial monotone boolean function and $U = f^{-1}(1)$, $V = f^{-1}(0)$, then $S^+_f$ is the $|U| \times |V|$ matrix defined by

$$S^+_f(x,y) = \sum_{i=1}^n [x_i = 1 \land y_i = 0]$$

for all $x \in U, y \in V$. To construct our embedding we create two mappings $\mu : \{0,1\}^m \rightarrow \text{Gen}_n^{-1}(1)$ and $\rho : \{0,1\}^m \rightarrow \text{Gen}_n^{-1}(0)$ such that $S^+_{\text{Gen}_n}(\mu(x),\rho(y)) = (\text{viol}_{\text{PEG}} \circ g^m)(x,y)$.

Let $G$ be any DAG with $m$ edges, and assume without loss of generality that $G$ has a single distinguished source node $s^*$ with a single outgoing edge, and sink nodes $u_1, \ldots, u_t$. We define an auxiliary graph $G^{aux}$ as follows. The nodes of the auxiliary graph $G^{aux}$ are the edges of $G$. Then, if $(u,v), (v,w)$ are edges of $G$ that share a node $v$, we add a directed edge $((v,w),(u,v))$ to $G^{aux}$ (note the reverse in the edge direction). Note the sink node of $G^{aux}$ corresponds to the unique edge leaving the source node of $s^*$, and the source nodes of $G^{aux}$ correspond to the edges that enter sinks of $G$.

Letting $N = 2^k$, for each node in $G^{aux}$ we create $N$ points in the resulting GEN instance, and we index each such point as $e_x$ for each $x \in \{0,1\}^k$. We also create a designated start node $1$ and a designated target node $n$.

Now we describe the functions $\mu, \rho$ in the embedding.

- **Definition of $\mu : \{0,1\}^m \rightarrow \text{Gen}_n^{-1}(1)$**. Let $x \in \{0,1\}^m$ and write $x = x_1 x_2 \cdots x_m$ where $x_i \in \{0,1\}^k$. Order the edges $e^1, e^2, \ldots, e^m$, and we regard $e^i$ as the unique edge leaving the distinguished source node. We regard each $x_i \in \{0,1\}^k$ as selecting the point $e^i_{x_i}$ from the resulting GEN instance; we then plant a copy of $G^{aux}$ inside the GEN instance on these points. That is, whenever a node $u = e^i$ in $G^{aux}$ has two incoming edges from $v = e^j, w = e^k$, we plant the triple $(e^j_{x_j}, e^k_{x_k}, e^i_{x_i})$; when $u$ has only a single incoming edge from $v$ we plant the triple $(e^j_{x_j}, e^i_{x_i}, e^i_{x_i})$. Finally, for each of source nodes $u = e^i$ of $G^{aux}$ we add the triple $(1,1,e^i_{x_i})$. It is easy to see that the result is in $\text{Gen}_n^{-1}(1)$, since we have explicitly planted the graph $G^{aux}$ connecting the source node $1$ to the target node $n$.

- **Definition of $\rho : \{0,1\}^m \rightarrow \text{Gen}_n^{-1}(0)$**. Given $y \in \{0,1\}^m$ we similarly write $y = y_1 y_2 \cdots y_m$ and add triples to the instance as follows. For any constraint of PEBG enforcing that a node $v$ in $G$ is not a proper sink (for example, by preventing the case where $e^i = (u,v)$ is in $G$ and $e^j = (v,w_1), e^k = (v,w_2)$ are not in $G$), and for any $x^i, x^j, x^k \in \{0,1\}^k$, we add the triple $(e^j_{x_j}, e^k_{x_k}, e^i_{x_i})$ to the instance iff the constraint is satisfied by the assignment $e^i = g(x_i, y_i), e^j = g(x_j, y_j), e^k = g(x_k, y_k)$. We apply a similar construction for each of the constraints corresponding to the source node of $s^*$ having an outgoing edge and set all other triples to 0. This is a 0-instance of $\text{Gen}_n$ since if it was a 1-instance it must contain an embedding of $G^{aux}$ (as in the construction of $\mu$ above), but this would imply that the corresponding assignment $x$ embedding $G^{aux}$ would satisfy every constraint of PEBG, which is a contradiction.
Finally, to see that $S_{\text{Gen}}^+(\mu(x), \mu(y)) = \text{viol}_{\text{Gen}}(g^m(x, y))$ observe that, by construction, each triple in $\mu(x)$ which does not occur in $\mu(y)$ corresponds exactly to a unique violated constraint of $\text{PEB}_G$. ▼

References


On Semi-Algebraic Proofs and Algorithms


42 Claire Mathieu and Alistair Sinclair. Sherali-adams relaxations of the matching polytope. In Michael Mitzenmacher, editor, Proceedings of the 41st Annual ACM Symposium on Theory of