Simulation of Arithmetical Circuits by Branching Programs with Preservation of Constant Width and Syntactic Multilinearity.

Maurice Jansen* and B.V. Raghavendra Rao**

Abstract. We study structural properties of restricted width arithmetical circuits. It is shown that syntactically multilinear arithmetical circuits of *constant* width can be efficiently simulated by syntactically multilinear algebraic branching programs of *constant* width, i.e. that sm-VSC^0 $\subseteq$ sm-VBWBP. Also, we obtain a direct characteriztion of *poly*-size arithmetical formulas in terms of *multiplicatively disjoint* constant width circuits, i.e. MD-VSC^0 = VNC^1.
For *log*-width *weakly-skew* circuits a width efficient multilinearity preserving simulation by algebraic branching programs is given, i.e. weaklyskew-sm-VSC^1 $\subseteq$ sm-VBP[width=$\log^2 n$].
Finally, coefficient functions are considered, and closure properties are observed for sm-VSC^i, and in general for a variety of other syntactic multilinear restrictions of algebraic complexity classes.

1 Introduction

In this paper the computational power of space bounded computation is studied in the arithmetical setting by considering arithmetical circuits of restricted width. For such circuits several elementary questions are still left unanswered. We are interested in the following question posed in [MR08]: can arithmetical circuits of constant width and polynomial degree be simulated by polynomial size arithmetical formulas ? If indeed so, this would yield an equivalence, since a simulation the other way around follows from a construction by Ben Or and Cleve [BOC92].

One strategy to approach the above question is to investigate under what additional assumptions one can indeed do the simulation. Mahajan and Rao show that every *syntactically multilinear* constant width circuit has an equivalent polynomial size arithmetical formula [MR08]. However, it was left open whether this arithmetical formula can be guaranteed to be syntactically multilinear. The starting point of this note is the observation that, with a careful modification, this can in fact be achieved. In other words, letting sm-VSC^0 and sm-VNC^1 stand for the syntactically multilinear restrictions of the arithmetical complexity classes corresponding to *poly*-size constant width circuits and *poly*-size circuits of $O(\log n)$-depth, respectively, we have the following theorem:

Theorem 1. sm-VSC^0 $\subseteq$ sm-VNC^1.

* Corresponding Author: Centre for Theory in Natural Science, Aarhus University, IT-Parken, Aabogade 34, DK-8200 Aarhus N, Denmark, phone: +45 8942 5600, mjjansen@daimi.au.dk.
** The Institute of Mathematical Sciences, Chennai, 600 113, India, bvrr@imsc.res.in.

Details regarding this modification will appear in the full version of this paper. Theorem 1 raises the following question: is it that syntactically multilinear constant width arithmetical circuits are equivalent to constant width syntactically multilinear branching programs? The main result of this paper is an answer to this question in the affirmative. Letting sm-VBWBP stand for the class corresponding to *poly*-size algebraic branching program of constant width, we strenghten Theorem 1 as follows:

Theorem 2. sm-VSC$^0 \subseteq$ sm-VBWBP.

Syntactically multilinear circuits can be assumed without loss of generality to be *multiplicatively disjoint* (See [MP06]). Hence it is natural to consider enhancements to the construction underlying Theorem 2 under assumption of multiplicative disjoinness, and also the even more stringent condition of *weak skewness* (See [MP06]). The best dependence on width we obtain under the latter restriction, allowing us to conclude $O(\log n)$-width, *poly*-size weakly-skew arithmetical circuits can be simulated by $O(\log^2 n)$-width, *poly*-size algebraic branching program with preservation of syntactic multilinearity. This will follow from a careful modification of the construction given in [Jan08].

In the general world, we observe that if a given constant width circuit is multiplicatively disjoint, then it can be depth reduced to yield a formula. To the best of our knowledge, this is the largest chunk of the class of constant width arithmetical circuits which are known to have equivalent formulas.

Raz established an $n^{\Omega(\log n)}$ lower bound for the syntactically multilinear formula size of the permanent [Raz04], which by our main theorem is now also established for constant width syntactically multilinear arithmetical circuits. However, the best known multilinear formula for the permanent is given by Reyser, which is of size $O(n2^n)$ [Rys63]. Raz and Yehudayoff strenghten the lower bound to $2^{n^{\Omega(1/d)}}$, for constant depth syntactically multilinear formulas with product depth d [RY08]. The latter can be simulated within sm-VBWBP, and this class appears to be the appropriate next place in line to strenghten the $n^{\Omega(\log n)}$ bound for the permanent. We interpret Theorem 2 as an obstacle to overcome in this, i.e. in spite of their conceptual simplicity, syntactically multilinear bounded width branching programs turn out to be more powerful than perhaps expected.

Following [Mal07], we study the complexity of coefficient functions of polynomials. Closure properties will be observed that hold quite universally among the syntactically multilinear circuit classes, in particular for the restricted width classes mentioned above. Although coefficient functions are known to be VNP hard even in the case of depth three arithmetic formula, generally syntactically multilinear circuit classes are closed for coefficient functions. Also, we show that if any coefficient function of a polynomial is in a syntactic multilinear class, then so is the polynomial itself. Hence, in the terminology of [Mal07], generally a syntactically multilinear arithmetic circuit class is *stable* for coefficient functions.

The rest of this paper is divided as follows. In Section 2 we introduce definitions and notations. Section 3 contains the construction for weakly-skew circuits. Section 4 contains a proof of Main Theorem 2. In Section 5 we study coefficient functions. Finally, we end in Section 6 by posing several open problems.

2 Preliminaries

For integer n, $[n]$ denotes the set $\{1, 2, \ldots, n\}$. Let R be a commutative ring with multiplicative identity 1. An *arithmetical circuit* C over R is a directed acyclic graph, with nodes (called *gates*) with labels taken from $\{+, \times\} \cup X \cup R$, where $X = \{x_1, x_2, \ldots, x_n\}$. A node with in-degree zero must take its label from X or R, depending on whether it represents and *input* or *constant gate*. C has at least one node of out-degree zero called an *output gate*. Every gate in g computes a polynomial in $R[X]$ defined in the usual way. The polynomial computed by the circuit is the polynomial computed by the output gate, and if C has more than one output gate, then this is the union of all such polynomials. For a gate f, $Var(f)$ denotes the subset of X of variables that appear in the subcircuit rooted at f. A circuit is *layered*, if the node set can be partitioned into a sequence of sets, called *layers*, with edges only between consecutive layers.

Fan-in (*fan-out*) of C is the maximum in-degree (out-degree) of any gate in C. The size of C is defined as the number of gates and edges (called *wires*) in C. If C has a constant fan-in then we mean size to be simply the total number of gates. *Depth* of a circuit is the length of longest directed path in the underlying graph. *Width* of a layered circuit is the maximum number of nodes at any layer. A *formula* is a circuit, where fan-out of every gate is bounded by one. The *formal degree* of a gate (degree for short) is defined inductively as follows: input gates have degree one, and for an addition or multiplication gate it is the sum or product of the degree of its inputs, respectively. The degree of a gate and the degree of the polynomial computed at the gate can differ due to cancellations.

A circuit C is said to be *skew*, if for every multilplication gate $f = g \times h$, one of g or h is an input or constant gate. C is said to be *weakly skew*, if for every $f = g \times h$, either the edge (g, f) or (h, f) is a bridge in the circuit, i.e removing the edge increase the number of weakly connected components. Further, in a *multiplicatively disjoint* (MD for short) circuit, for every gate $f = g \times h$, the sub-circuits rooted at g and h are disjoint (as graphs). Note that MD-circuits are generalisations of weakly skew circuits, which in turn contains skew circuits.

The following are the arithmetical circuit classes that will be used in the next sections. These contain families of polynomials $(f_m)_{m\geq 1}$, where $f_m \in R[x_1, x_2, \ldots, x_{p(m)}]$, for some polynomial p. The measures size, width, depth, fan-in and fan-out are all defined in terms of m.

VAC^0	: *poly*-size, constant-depth, unbounded fan-in circuits.
VNC^1	: *poly*-size, *log*-depth, constant fan-in circuits.
VF	: *poly* size formulas.
VsSC^i	: *poly*-size, $O(\log^i m)$-width, *poly*-degree circuits.
VSAC^1	: *poly*-size, *log*-depth, constant $\times$-fan-in circuits.
VP	: *poly*-size, *poly*-degree circuits.

The class VNP is defined as families of polynomials $(g_m)_{m\geq 1}$, where $g_m = \sum_{b\in\{0,1\}^{q(m)}} f_m(x, b)$, for polynomial $q(m)$ and $(f_m)_{m\geq 1} \in \mathrm{VP}$.

An *algebraic branching program* (BP for short) is a directed acyclic graph, where edges are labeled from $X \cup R$. There are two designated nodes, s and t, where s has in-degree zero and t has out-degree zero. We assume that a BP is layered. Size of a BP is the number of nodes and edges in it and width is the maximum number of nodes at any layer. Length of a BP is the number of layers in it. Depth of a BP B equals 1+length(B). The polynomial computed by a BP P, is the sum of weights of all s-t paths in P, where the weight of a path is the product of all edge labels in the path. We will also consider multi output BPs, where the above is generalized in the obvious way to several nodes $t_1, t_2, \ldots, t_m$ existing at the last level. Note that BPs can be simulated by skew circuits and vice versa with a constant blow up in the width. We let VBP, VLWBP and VBWBP stand for classes corresponding to *poly*-size BPs of *poly*, *log* and *constant* width, respectively. Known relationships among the classes defined so far are: $\text{VAC}^0 \subseteq \text{VNC}^1 = \text{VF} = \text{VBWBP}$ [Bre73,BOC92], and $\text{VBP} \subseteq \text{VSAC}^1 = \text{VP} \subseteq \text{VNP}$ [VSBR83].

A circuit C is said to be *syntactic multilinear*, if for every multiplication gate $f = g \times h$ in C, $Var(g) \cap Var(h) = \emptyset$. A BP is said to be syntactic multilinear, if on every s-t path every variable appears at most once. For arithmetical circuit classes we add the prefix 'sm-' to indicate the syntactic multilinear version of the corresponding class. For the class sSC^i we drop one 's' as *poly* bounded degree is implied by syntactic multilinearity, i.e. we write sm-VSC^i and also MD-SC^i. Known relationships are: $\text{sm-VBWBP} \subseteq \text{sm-VsSC}^0 \subseteq \text{sm-VNC}^1 = \text{sm-VF}$ ([MR08]) and $\text{sm-VBP} \subseteq \text{sm-VSAC}^1 = \text{sm-VP}$ ([RY08]).

The above results together with the main theorem of our paper lead to the scenario as depicted in Figures 1 and 2. The main contrasting point is that $\text{VBWBP} = \text{VNC}^1 \subseteq \text{VsSC}^0$, whereas $\text{sm-VSC}^0 = \text{sm-VBWBP} \subseteq \text{sm-VNC}^1$.

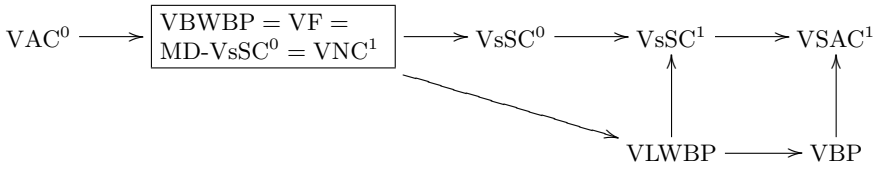

Fig. 1. Relationship among arithmetic classes around arithmetic formulas

3 Weakly Skew Circuits

The main objective of this section is to prove the following result:

Theorem 3. weaklyskew-sm-VSC^0 = sm-VBWBP.

Lemma 1. *Let C be an arithmetical circuit of width w and size s. Then there is an equivalent arithmetic circuit C' of width $O(w)$ and size poly(s) such that fan-in and fan-out of every gate is bounded by two, and every layer has at most one*

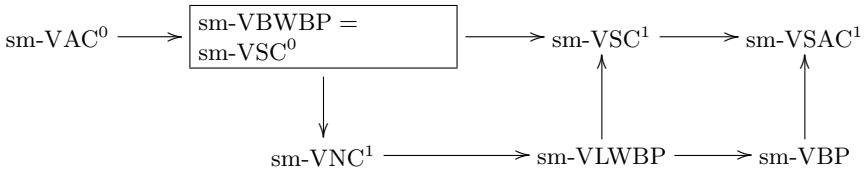

Fig. 2. Syntactic Multilinear Versions

$\times$ gate. Moreover, C' preserves any of the properties of syntactic multilinearity, weakly-skewness and multiplicatively disjointness.

Proof. Let k be the bound on maximum fan-in and fan-out of C. First we can reduce the fan-in to two by staggering the circuit and keeping copies of the gates as and when needed. This blows up the width to $2w$ and size to wks. Now in a similar manner we can ensure that the fan-out of a gate is bounded by two and the size blow up will now be w^2k^2s and width will be $4w$. To ensure the second condition we need to push the gates (using staggering and dummy + gates) up by at most $4w$ levels, thus making the total width $8w$ and size $2w^2k^2s$. Since $k \leq w + n$ and $w \leq s$ we have size bounded by $poly(s, n)$. □

For a BP B of depth d with a single source s, we say B *is endowed with a mainline*, if there exist nodes $v_1, v_2, \ldots, v_{d-1}$ reachable only along the path $s, v_1, v_2, \ldots, v_{d-1}$, and if the labels on this path are all set to the field constant 1. We require this feature to ensure our construction builds programs with a single source. For BPs B_1 and B_2, *piping* the mainline of B_1 into the mainline of B_2 is the operation of removing the edge from the source of B_2 to the first node v of the mainline of B_2, and adding an edge from the last node w of the mainline of B_1 to v.

In the following, as a typical next step we draw an edge from an output of B_1 to the source of B_2. This can be from last node w, if we just want to compute B_2 or some other output in case we want to do multiplication. Note the width of this composition is bounded by $\max(\text{width}(B_1), \text{width}(B_2))$.

Lemma 2. *Let C be a weakly skew circuit of width $w > 1$ and size $s > 1$ as given by Lemma 1. Let $f_1, \ldots, f_w$ be the output gates of C. Then there exists an equivalent arithmetical BP $[C]$ of width $w^2 + 1$, length $O(2^w s)$ and size $O((w^2 + 1)2^w s)$. $[C]$ has a single start node b and and terminal nodes $[f_1], \ldots, [f_w], v$ and will be endowed with a mainline ending in v. Moreover, if C is syntactically multilinear then so is $[C]$.*

Proof. We proceed by induction on both s and w. If $s = 2$, the lemma holds trivially. If $w = 2$, C is a skew-circuit hence can be seen as a BP of width 3 (We also need to add a mainline hence width is 3).

Let $s > 2$ and $w > 2$ be given, and assume that C has at least 2 layers. By induction hypothesis, the lemma holds for all circuits of size s' and w', where either $s' < s$ and $w' \leq w$ or $s' \leq s$ and $w' < w$.

Wlog. assume that f_1 is a $\times$ gate and $f_2, \ldots, f_w$ are $+$ gates. Let C' be the circuit obtained by removing the output gates of C. Let $g_1, \ldots, g_w$ be the output gates of C'. Wlog. assume $f_1 = g_1 \times g_2$, and also that the edge (g_1, f_1) is a bridge in the circuit. We define subcircuits D, E and of C' as follows: D is obtained from C' by deleting the sub-circuit rooted at g_1, E is the sub-circuit rooted at g_1. Let $s' = size(C'), w' = \text{width}(C')$, $s_J = \text{size}(J)$ and $w_J = \text{width}(J)$ for $J \in \{D, E\}$. Note that $s = s' + w$, and $s_J < s$ for $J \in \{D, E\}$.

By induction hypothesis, we have branching programs $[D]$ and $[E]$, both endowed with a mainline. Let $[g_1], v'$ denote the output of E and $[g_2], \ldots, [g_w], v''$ denote the output nodes of $[D]$, where v' and v'' are the last nodes on the mainlines. Let $[F]$ be the subprogram of $[D]$, which consists of all paths from the source of $[D]$ to $[g_2]$ and v''. Construct the program $[C]$ with output nodes $[f_1], \ldots, [f_w], v$ as follows:

case 1: $w_E \leq w - 1$.

We compose $[D]$ before $[E]$ as follows:

1. For $i, j \geq 2$, $[g_j]$ has an edge to $[f_i]$ iff g_j is an input to f_i.
2. For input gates f_i, draw an edge from v'' to $[f_i]$ with the appropriate label.
3. Add an edge from $[g_2]$ to $[f_1]$.
4. Identify the start node of $[E]$ with $[f_1]$ and relabel the output node of $[E]$ as $[f_1]$. Pipe the mainline of $[D]$ into the mainline of $[E]$.
5. Stagger the nodes $[f_2], \ldots, [f_w]$ until the last level of the new program.

Size and width analysis: By induction hypothesis $\text{width}([E]) \leq (w_E)^2 + 1 \leq (w-1)^2 + 1$, and $\text{width}([D]) \leq w^2 + 1$, and $\text{length}([E]) \leq 2^{w-1} size(E)$ and $\text{length}([D]) \leq 2^w size(D)$. Now $\text{width}([C]) = \max\{\text{width}([D]), \text{width}([E]) + w - 1\} \leq w^2 + 1$ and $\text{length}([C]) = \text{length}([D]) + \text{length}([E)] \leq 2^{w_D} s_D + 2^{w_E} s_E \leq 2^w s_D + 2^{w-1} s_E \leq 2^w s$ as $s = s_D + s_E + w$.

case 2: $w_E = w$, and hence $w_F \leq w - 1$ and $w_D \leq w - 1$.

We compose $[E]$ before $[F]$ before $[D]$ as follows:

1. Identify $[g_1]$ with the source of $[F]$, and pipe the mainline of $[E]$ into the mainline of $[F]$.
2. Add an edge from v' (last node of mainline of $[F]$) to the source of $[D]$,
3. Pipe the mainline of $[F]$ into the mainline of $[D]$.
4. Alongside $[D]$ stagger the output of $[F]$ (which equals $[f_1]$).
5. For $i, j \geq 2$, $[g_j]$ has an edge to $[f_i]$ iff g_j is an input to f_i.
6. Finally, for input gates f_i, draw an edge $(v'', [f_i])$ with the appropriate label.

Size and width analysis: By induction hypothesis, $\text{width}([E]) \leq w^2 + 1$, $\text{width}([D]) \leq (w-1)^2 + 1$. Hence also $\text{width}([F]) \leq (w-1)^2 + 1$. Observe, $\text{width}([C]) \leq max(\text{width}([E]), \text{width}([F]), \text{width}([D]) + 1) \leq w^2 + 1$.

Now, $\text{length}([C]) = \text{length}([E]) + \text{length}([F]) + \text{length}([D]) + 1 \leq 2^w s_E + 2^{w-1} s_F + 2^{w-1} s_D + 1 \leq 2^w (s_D + s_E) + 1 \leq 2^w s$. Since size of a layered BP is length $\times$ width we have the required size bound. If C was syntactic multilinear to start with, then it is easy to see that so is $[C]$.

□

Corollary 1. weaklyskew-VSC^0 = VNC^1 = VBWBP.

Corollary 2. weaklyskew-sm-VSC$^1 \subseteq$ sm-VBP[width = $\log^2 n$].

Conversion to Weakly Skew: We note that it is possible to process a multiplicatively disjoint circuit into a weakly-skew circuit with preservation of syntactic multilinearity.

Lemma 3. *For any leveled syntactically multilinear multiplicatively disjoint circuit C of width $w \geq 1$ and size $s \geq 1$ such that each layer has at most one multiplication gate, there exists a leveled syntactically multilinear weakly-skew circuit $[C]$ of size at most s^w such that for any gate g of C, there is a gate $[g]$ in $[C]$ that computes the same polynomial.*

Proof. We prove the above lemma by induction on both s and w. We have two base cases: $s = 1$ and $w = 1$. In both cases the lemma trivially holds.

Let $s > 2$ and $w > 1$ be given. By induction hypothesis, the lemma holds for any circuit of size s' and width w' for which either $s' < s$ or $w' < w$.

Case I: The output layer has addition and input gates only.

Let C' be the circuit C excluding the output layer. Recursively process C' and add the output layer back to form $[C]$ from $[C']$. We have that size($[C]$) $\leq$ size($[C']$)$^w + w \leq (s - w)^w + w \leq s^w$.

Case II: The output layer contains a multiplication gate f.

Let C_1 and C_2 be the subcircuits rooted at the inputs of the multiplication gate f. Since C_1 and C_2 are disjoint, one of them is guaranteed to have width at most $w - 1$. Wlog. assume C_1 has width at most $w - 1$. Let $g_1, g_2, \ldots, g_m$ be all gates not in C_1 that take input from C_1. Let D be the subcircuits formed by the gates in C excluding C_1, where any input g taken by a g_i from C_1 is removed. Let s_d = size(D) and s_1 = size(C_1). Then $s = s_d + s_1$. Recursively process D and C_1 (separately) to obtain weakly skew circuits $[D]$ and $[C_1]$ of sizes s_d^w and s_1^{w-1}, respectively. Now we put back removed inputs to each of $[g_1], [g_2], \ldots, [g_m]$ from the appropriate gate in $[C_1]$.

The circuit we obtain from $[C_1]$ and $[D]$ this way is almost weakly skew. The only issue is that the adding back of original inputs from say $[g] \in [C_1]$ at input $[g_i]$ can violate the weak skewness condition for $[g_i]$ and also for gates in $[C_1]$. We resolve this by simply duplicating the subcircuit rooted at $[g]$. Observe that size($[C]$) $\leq s_d^w + m \cdot s_1^{w-1} \leq s_d^w + s_d \cdot s_1^{w-1} \leq (s_d + s_1)^w = s^w$.

□

The above gives an alternative proof of sm-VSC$^0 \subseteq$ sm-VBP. Namely, by above lemma we get that sm-VSC$^0 \subseteq$ weaklyskew-sm-VP. Next use the construction from [Jan08] that shows weaklyskew-sm-VP $\subseteq$ sm-VBP. The other way to arrive at this is by Theorem 1. However, there the size of the resulting BP is $O(2^{w^2} s^{25w})$. In this regard, Lemma 3 still yields a slightly better size bound than the construction underlying Theorem 2, since there the resulting size is $O(w^2 s^w)$.

4 Multiplicatively Disjoint Circuits

In this section we prove Theorem 2. In fact, we prove that multiplicatively disjoint circuits of constant width and polynomial size can be simulated by BPs of constant width and polynomial size preserving the syntactic multilinearity property. In general multiplicatively disjoint circuits are equivalent to polynomial degree circuits (see [MP06]). The following theorem can be deduced from [MR08]:

Theorem 4. MD-VSC^0 = VNC^1

Proof. Let C be a multiplicatively disjoint arithmetic circuit of width w and size s. Let $X = \{x_1, x_2, \ldots, x_n\}$ be the set of variables in the circuit. Construct a new circuit by replacing j^{th} occurence of x_i by a new variable $y_{i,j}$, for all $i.j$. Note that C' is a circuit with at most ns many variables and of size s and width w. Also, as C is multiplicatively disjoint, C' is syntactic multilinear in the variables $Y = \{y_{1,1}, y_{1,2} \ldots, y_{n,s}\}$. Now applying the construction of [MR08], we get an arithmetic formula of depth $O(w^2 \log s)$ and size $O(2^{w^2} s^{25w})$, but as w a constant, we get the required formula by replacing the $y_{i,j}$s with x_is. Now the equivalence follows from [BOC92]. □

Remark 1. Note that the above theorem does not already give Theorem 2, as the only known procedure to convert an arithmetic formula into an equivalent bounded width branching program of [BOC92] does not preserve syntactic multilinearity (For an example see [MR08]).

We strengthen the above result by giving a direct construction of BPs from multiplicatively disjoint circuits.

Lemma 4. *C be a multiplicatively disjoint arithmetical circuit of width w and size s as given by Lemma 1. Let $f_1, \ldots, f_w$ be the output gates of C. Then there exists an equivalent arithmetical branching program $[C]$ of width $O(w^2)$, length $O(s^w)$, and size $O(w^2 s^w)$. $[C]$ has a single start node b and and terminal nodes $[f_1], \ldots, [f_w], v$, and will be endowed with a mainline ending in v. Moreover, if C is syntactic multilinear then so is $[C]$.*

Proof. The proof is similar to that of Lemma 2. We proceed by induction on both s and w. If $s = 2$, the lemma holds trivially. If $w = 1$, C is a skew-circuit, and hence can be seen as a BP of width 3 (by adding a mainline).

Let $s > 2$ and $w > 2$ be given, and assume that C has at least 2 layers. Suppose, by induction hypothesis that the lemma holds for all circuits of size s' and w', where either $s' < s$ and $w' \leq w$ or $s' \leq s$ and $w' < w$.

Let C' be the sub-circuit obtained by deleting $f_1, \ldots, f_w$. Let $G = \{g_1, \ldots, g_w\}$ be the output gates of C'. Wlog. let $f_1 = g_1 \times g_2$ be the only multiplication gate at the output layer of C. Let D denote the sub-circuit rooted at g_1. Since C is multiplicatively disjoint, we have either width$(D) \leq w - 1$ or width$(E) \leq w - 1$. Wlog. assume that width$(D) \leq w - 1$.

Let $s' = \text{size}(C')$, $s_D = \text{size}(D)$, $w' = \text{width}(C')$, and $w_D = \text{width}(D)$. By induction hypothesis, we obtain BPs $[C']$ and $[D]$. $[C']$ has $w+1$ output nodes, namely $[g_1], \ldots, [g_w], v$. $[D]$ has two output nodes $[g'_1]$ and v'.

Now construct the BP $[C]$ with output nodes $[f_1], \ldots, [f_w], v$ by composing $[C']$ before $[D]$ as follows: For all $i \geq 2$, connect $[g_j]s$ to $[f_i]s$ according the edges in the circuit C, i.e edge $([g_j], [f_i])$ is in $[C]$ iff g_j is an input for f_i. In case f_i is an input gate, draw an appropriately labeled edge from v. Put an edge from $[g_2]$ to $[f_1]$. Now identify the start node of $[D]$ with $[f_1]$ and re-label the terminal node of $[D]$ as $[f_1]$. Do the necessary staggerings to carry on the values $f_2, \ldots, f_w$ to the last layer. We also pipe the mainline of $[C']$ into the mainline of $[D]$.

Analysis: By induction hypothesis, we have $\text{length}([C']) \leq s'^{w'} \leq (s-w)^w$ as $s' = s - w$ and $w' \leq w$. Furthermore, $\text{width}([C']) \leq w'^2 + 1 \leq w^2 + 1$, $\text{length}([D]) \leq s_D^{w_D} \leq (s-w)^{w-1}$, and $\text{width}([D]) \leq (w-1)^2 + 1$ as $s_D \leq s - w$ and $w_D \leq w - 1$.

Now by the construction, $\text{width}([C]) = \max\{\text{width}([C'], \text{width}([D]) + w - 1\} \leq \max\{w^2+1, (w-1)^2+w-1\} \leq w^2+1$. Hence, $\text{length}([C]) = \text{length}([C']) + \text{length}([D]) \leq (s-w)^w + (s-w)^{w-1} \leq s^w$, for $w > 2$ and $w < s$. Hence $\text{size}([C]) = (w^2+1)s^w$. It is easy to see that this construction preserves the syntactic multilinearity property. □

4.1 Proof of Main Theorem 2

Given a syntactically multilinear circuit C of width w and size s, we first replace all the wires carrying only ring constants in C by new variables (as done in [MR08]), to get a circuit D of width $w_d \leq w^2$ and size $s_d \leq ws$. Note that the circuit D is multiplicatively disjoint. By Lemma 4 we get a syntactic multilinear BP $[D]$ of width $w_d^2 + 1$ and size s_d^w. Now replacing the introduced variables by the original constants they represent, we get the required syntactic multilinear BP $[C]$. □

Remark 2. By closer inspection of how Lemma 4 deals with input gates one can actually conclude $\text{width}([D]) \leq w^2 + 1$ and $\text{size}([D]) \leq O(s^w)$ in the above.

5 Coefficient Functions

Let f be a polynomial over variables $X = \{x_1, x_2, \ldots, x_n\}$. For a monomial m in variables from X, the partial coefficient function $coef(f, m)$ is defined to be the unique polynomial g for which there exists polynomial h with none of its monomials divisble by m such that $f = mg + h$.

Malod studies the complexity of computing coefficient functions computed by class of arithmetic circuits [Mal07]. From an old observation by Hammon, it can be seen that the permanent polynomial equals $coef(f, y_1 y_2 \ldots y_n)$, where f is given by the following depth three formula $f = \prod_{i \in [n]} \sum_{j \in [n]} x_{ij} y_j$. In [Mal07] it is shown that Hamiltonian polynomial can be represented as a coefficient of a polynomial g computed by polynomial size arithmetic circuits. A closer

inspection shows that this g is actually in VBP. Thus the arithmetic circuit classes which are at least as powerful as VAC^0 can generate VNP complete polynomials as coefficient functions, and hence the coefficient functions in general are hard.

In the case of polynomials computed by syntactic multilinear circuits we will prove that the situation is markedly different compared to the general case. For a multilinear polynomial f over variables $x_1, x_2, \ldots, x_n$, we define coefficient function $mcoef(f, \cdot)$ by $mcoef(f, a_1, a_2, \ldots, a_n) = coef(f, x_1^{a_1} x_2^{a_2} \ldots x_n^{a_n})$, for any $a \in \{0,1\}^n$. Corresponding to $mcoef(f, \cdot)$ is a unique multilinear polynomial $g(x, e)$ in variables from X and E, such that for all $a \in \{0,1\}^n$, $g(x, a) = mcoef(f, a)$. Per abuse of notation we will denote this g by $mcoef(f, a)$.

5.1 Closure Property

A syntactically multilinear complexity class sm-V$\mathcal{C}$ is said to be *closed under taking coefficients*, if for any $f \in$ sm-V$\mathcal{C}$, $mcoef(f, e) \in$ sm-V$\mathcal{C}$. We have the following identities: for any polynomials f and g,

$$mcoef(f + g, e) = mcoef(f, e) + mcoef(g, e). \tag{1}$$

For polynomials f and g with $Var(f) \cap Var(g) = \emptyset$,

$$\begin{aligned} mcoef(fg, e) = {} & mcoef(f, e)[\ e_i := 0 \text{ for } x_i \notin Var(f)] \cdot \\ & mcoef(g, e)[e_i := 0 \text{ for } x_i \notin Var(g)] \cdot \\ & \prod_{x_i \notin Var(f) \cup Var(g)} (1 - e_i) \end{aligned} \tag{2}$$

For individual variables and constants μ we have $mcoef(x_i) = (x_i(1 - e_i) + e_i) \prod_{j \in [n], j \neq i} (1 - e_j)$, and $mcoef(\mu) = \mu \prod_{j \in [n]} (1 - e_j)$.

Theorem 5. *Each of the following syntactically multilinear classes is closed under taking coefficients:* sm-VP, sm-VBP, sm-VNC^1, sm-VSC^i, sm-VBWBP*, and* sm-VAC^i*, for all* $i \geq 0$.

Proofsketch. For formula classes sm-VNC^1 and AC^i it is immediately clear how to convert a formula Φ computing f into a formula computing $mcoef(f, e)$ using Equations (1) and (2). For circuits one has to ensure the substitution at a multiplication gate $g = g_1 \times g_2$ using Equation (2) are consistent with other uses of g_1 and g_2. This can be guaranteed by first levelling the circuit with alternating layers of multiplication and addition gates. □

Consequently, we have no analogue of Hammon's observation for the permanent with $f \in$ sm-VNC$_1$ by [Raz04].

Corollary 3. *Permanent (and also Determinant) cannot be expressed as a coefficient of some monomial of a polynomial computed by a syntactically multilinear arithmetic formula of polynomial size.*

5.2 Stability

Following [Mal07], we say a complexity class sm-V$\mathcal{C}$ is *stable for coefficient functions*, if sm-V$\mathcal{C}$ is closed under taking coefficients and whenever $mcoef(f, e) \in$ sm-V$\mathcal{C}$, then $f \in$ sm-V$\mathcal{C}$. For a multilinear polynomial $f(x, e)$, let $\Sigma(E)\ f$ denote $\sum_{b\in\{0,1\}^m} f(x, b)$. We say a complexity class sm-V$\mathcal{C}$ is *closed under taking exponential sums*, if whenever $f(x, e) \in$ sm-V$\mathcal{C}$, then $\Sigma(E)f \in$ sm-V$\mathcal{C}$. Again the situation is contrary to the non-multilinear case, e.g. one can obtain the permanent as $\Sigma(E)\ f$, for $f \in \mathrm{VNC}^1$ [Val82], cf. [B$\ddot{0}$0].

Theorem 6. *The following are closed under exponential sums, and hence stable for coefficient functions:* sm-VP, sm-VBP, sm-VNC^1, sm-VSC^i, sm-VBWBP, *and* sm-VAC^i, *for all* $i \geq 0$.

The theorem will follow from the following straightforward proposition by patching a given circuit at gates with constant multiplications of appropriate powers of two. A proof will appear in the full version of the paper.

Proposition 1. *Let f and g be multilinear polynomials over X and E. We have that*

$$\Sigma(E)\ (f + g) = 2^a \Sigma(\mathit{Var}(f) \cap E)\ f + 2^b \Sigma(\mathit{Var}(g) \cap E)\ g,$$

where $a = |E - \mathit{Var}(f)|$ and $b = |E - \mathit{Var}(g)|$. Furthermore, if f and g are defined on disjoint variables sets,

$$\Sigma(E)\ fg = 2^c \Sigma(\mathit{Var}(f) \cap E)\ f \cdot \Sigma(\mathit{Var}(g) \cap E)\ g,$$

where $c = m - |\mathit{Var}(f) \cup \mathit{Var}(g)|$.

6 Open Problems

We ask the following four questions:

- Is sm-VSC$^1 \subseteq$ sm-VBP ?
- Is weaklyskew-sm-VSC$^1 \subseteq$ LWBP?
- Is VsSC$^0 \subseteq$ VBP ?
- Can we preserve width in Theorem 1 of [MP06]? If so, VsSC^0 = VNC^1.

Acknowledgements The second author thanks Meena Mahajan for enlightening discussions and useful feedback.

References

[B$\ddot{0}$0] Peter. Bürgisser. *Completeness and Reduction in Algebraic Complexity Theory.* Algorithms and Computation in Mathematics. Springer-Verlag, 2000.

[BOC92] Michael Ben-Or and Richard Cleve. Computing algebraic formulas using a constant number of registers. *SIAM J. Comput.*, 21(1):54–58, 1992.

[Bre73] Richard P. Brent. The parallel evaluation of arithmetic expressions in logarithmic time. In *Proc. Symp., Carnegie-Mellon Univ., Pittsburgh, 1973*, pages 83–102. 1973.

[CMTV98] Hervé Caussinus, Pierre McKenzie, Denis Thérien, and Heribert Vollmer. Nondeterministic NC^1 computation. *Journal of Computer and System Sciences*, 57:200–212, 1998.

[Jan08] Maurice J. Jansen. Lower bounds for syntactically multilinear algebraic branching programs. In *MFCS*, pages 407–418, 2008.

[Mal07] Guillaume Malod. The complexity of polynomials and their coefficient functions. In *IEEE Conference on Computational Complexity*, pages 193–204, 2007.

[MP06] Guillaume Malod and Natacha Portier. Characterizing valiant's algebraic complexity classes. In *MFCS*, pages 704–716, 2006.

[MR08] Meena Mahajan and B. V. Raghavendra Rao. Arithmetic circuits, syntactic multilinearity, and the limitations of skew formulae. In *MFCS*, pages 455–466, 2008.

[Raz04] Ran Raz. Multi-linear formulas for permanent and determinant are of super-polynomial size. In *STOC*, pages 633–641, 2004.

[RY08] Ran Raz and Amir Yehudayoff. Lower bounds and separations for constant depth multilinear circuits. In *CCC '08*, pages 128–139, 2008.

[Rys63] H. J. Ryser. Combinatorial mathematics. In *Carus Mathematical Monograph 14*, page 23. Mathematical Association of America, Washington DC., 1963.

[Val82] Leslie G. Valiant. Reducibility by algebraic projections. *L'Enseignement mathématique*, 28:253–268, 1982.

[VSBR83] Leslie G. Valiant, Sven Skyum, S. Berkowitz, and Charles Rackoff. Fast parallel computation of polynomials using few processors. *SIAM J. Comput.*, 12(4):641–644, 1983.