Multilinear Algebraic Branching Programs and the Min-Partition Rank Method

Théo Borém Fabris
University of Copenhagen, Denmark

Nutan Limaye
IT University of Copenhagen, Denmark

Srikanth Srinivasan
University of Copenhagen, Denmark

Amir Yehudayoff
University of Copenhagen, Denmark
Technion-ITT, Haifa, Israel

Abstract

It is a long-standing open problem in algebraic complexity to prove lower bounds against multilinear algebraic branching programs (mlABPs), however the best lower bounds are still quadratic (Alon, Kumar and Volk (*Combinatorica* 2020)). At the same time, it remains a possibility that the "min-partition rank" method introduced by Raz (*Theory Comput.* 2006), which is used to prove all known multilinear lower bounds, can also be used to prove superpolynomial lower bounds on the size of mlABPs. In this paper, we analyze the potential of the min-partition rank method to prove lower bounds on the size of mlABPs, and show the following results:

1. We relate this method to a purely combinatorial question regarding the minimum size of set systems whose chains satisfy a discrepancy condition. In the case of set-multilinear ABPs, this combinatorial measure *characterizes* the best lower bound that can be achieved via the min-partition rank method.
2. We prove a non-trivial upper bound on the size of a set system satisfying this combinatorial property. Together with our construction of full-rank mlABPs from set systems, this recovers a superpolynomial separation between mlABPs and multilinear formulas (Dvir, Malod, Perifel and Yehudayoff (*STOC 2012*)) via a conceptually different proof.
3. The property we study extends combinatorial notions of "balancing sets" considered in previous works, for which near-tight bounds are known via intervals families. We show that any intervals set system is very far from satisfying our property. This showcases how our methods capture combinatorial structures that evade previous techniques, and also allows us to improve and generalize known lower bounds for sum of ordered set-multilinear ABPs (Chatterjee, Kush, Saraf, Shpilka (*CCC 2024*)).

These results build a bridge between algebraic complexity theory and the behavior of random walks. Our upper bound uses the fact that, with noticeable probability, a random walk of length n on the integers returns to its starting point at least once every $n/\log n$ steps (Csáki, Erdős, and Révész (*PTRF 1985*)), while, for our lower bound, we prove that two independent random walks are "far" from each other in *discrete Fréchet distance*.

2012 ACM Subject Classification Theory of computation → Algebraic complexity theory

Keywords and phrases Algebraic branching programs, Multilinear computations, Rank methods

Digital Object Identifier 10.4230/LIPIcs.CCC.2026.22

Related Version *Full Version*: `https://eccc.weizmann.ac.il/report/2026/001/`

Funding *Théo Borém Fabris*: Supported by the European Research Council (ERC) under grant agreement no. 101125652 (ALBA).
Nutan Limaye: Supported by Independent Research Fund Denmark (grant agreement No. 10.46540/3103-00116B) and Basic Algorithms Research Copenhagen (BARC), funded by VILLUM Foundation Grant 54451.

41st Computational Complexity Conference (CCC 2026).
Editor: Dana Moshkovitz; Article No. 22; pp. 22:1–22:20

Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Srikanth Srinivasan: Supported by the European Research Council (ERC) under grant agreement no. 101125652 (ALBA).
Amir Yehudayoff: Supported by DNRF Chair grant no. 20231101-28697.

1 Introduction

Basic background. Algebraic complexity theory is the study of algebraic problems in models such as algebraic formulas, branching programs and circuits. Analogous to the Boolean circuit classes $\mathbf{NC}^1 \subseteq \mathbf{NL} \subseteq \mathbf{NC}^2$, in algebraic complexity we study complexity classes $\mathbf{VF} \subseteq \mathbf{VBP} \subseteq \mathbf{VP}$, and a central goal is to prove lower bounds against all classes in this chain.

Notable success has been achieved in the *multilinear* setting where we restrict each of the algebraic models to compute only multilinear polynomials (in a syntactic[1] way). In particular, this gives us a multilinear hierarchy $\mathbf{mlVF} \subseteq \mathbf{mlVBP} \subseteq \mathbf{mlVP}$. Well-known results of Raz [15, 14] showed superpolynomial lower bounds on the size of multilinear formulas computing some explicit polynomials in $\mathbf{mlVP}$, thus separating the first and third levels of this hierarchy. Subsequently, Dvir, Malod, Perifel and Yehudayoff [7] proved a stronger separation between the first and second levels of this hierarchy. There has been a large volume of follow-up work in algebraic complexity on extending the known techniques to different algebraic models (e.g. [16, 18, 19, 5, 2, 10, 11, 4]), and they also found applications in quantum computation [1, 19] and proof complexity [17, 9]. Despite this interest, we have not yet been able to extend our current lower bounds techniques to prove superpolynomial lower bounds for higher classes in the multilinear hierarchy such as $\mathbf{mlVBP}$. The main motivation for this work is to understand whether it is possible to extend known lower bound techniques from the multilinear setting to obtain superpolynomial lower bounds against $\mathbf{mlVBP}$.

Multilinear algebraic branching programs (mlABPs). The mlABP computational model is an algebraic analogue of Boolean branching programs, which model space-eficient (Boolean) computation. An mlABP is defined to be a directed acyclic graph G with a source a and sink b. The edges of G are labelled with affine linear polynomials (e.g. $1 - 2x_1 + x_2$). The polynomial computed is the sum, over all a-to-b paths, of the products of the edge-labels occurring along the path. To keep the polynomial multilinear, we impose the condition that the same variable cannot appear in more than one edge label along a path. Multilinear formulas and circuits are defined similarly using algebraic analogues of Boolean formulas and Boolean circuits, but we omit their formal definitions here as they are not crucial in what follows. The central open question we care about is the following:

▶ **Question 1.** Are there explicit n-variate multilinear polynomials P_n that cannot be computed by mlABPs of size $\mathrm{poly}(n)$?

As mentioned above, this question has been resolved positively for multilinear formulas by the work of Raz [15, 14]. For mlABPs and the stronger model of syntactially multilinear circuits, the best lower bounds are nearly quadratic [16, 2].

[1] We can also define a *semantic* notion of multilinearity, which coincides with the syntactic notion for formulas in terms of computational power, but is potentially more powerful in the setting of branching programs and circuits. In this paper, the word "multilinear" will only refer to the syntactic variants.

The lower bound method. Nearly all known lower bounds for multilinear models of computation (including all those mentioned above) can be proved using the "min-partition rank" technique developed by Raz [15][2], which relies on ideas of Nisan [12] and Nisan and Wigderson [13].

Given a multilinear polynomial $P(x_1, \ldots, x_n)$ with n even and a partition $\Pi = (Y, Z)$ of the underlying variable set into two parts, we define the *coefficient matrix* $M_\Pi(P)$ to be a $2^{|Y|} \times 2^{|Z|}$ matrix whose rows and columns are labelled by multilinear monomials m_Y and m_Z respectively. The (m_Y, m_Z)-th entry of $M_\Pi(P)$ is the coefficient of $m_Y m_Z$ in P. We define the *min-partition rank* of P by minimizing the rank of $M_\Pi(P)$ over all *equipartitions* (i.e. $|Y| = |Z| = n/2$):

$$\Gamma(P) := \min_{\substack{\Pi \\ \text{an equipartition}}} \operatorname{rank}(M_\Pi(P)).$$

We call the polynomial P *full-rank* if $\Gamma(P) = 2^{n/2}$. In other words, M_Π has full rank for all Π.

Using this notion of the complexity of a polynomial, Raz established the following.

▶ **Theorem 1** (Raz [15, 14]). *Let $\mathbb{F}$ be any field. Any multilinear formula computing a full-rank multilinear polynomial $P(x_1, \ldots, x_n) \in \mathbb{F}[x_1, \ldots, x_n]$ must have size at least $n^{\Omega(\log n)}$.*

Further, if $\mathbb{F}$ is infinite, then there exist multilinear circuits of size $\operatorname{poly}(n)$, and multilinear ABPs and formulas of size at most $n^{O(\log n)}$ computing full-rank polynomials.[3]

Using the above theorem, Raz proved the existence of explicit sequences of n-variate polynomials that have small multilinear circuits but no multilinear formulas of size $\leq n^{o(\log n)}$. A follow-up result of Dvir, Malod, Perifel and Yehudayoff [7] exhibited a similar separation between mlABPs and multilinear formulas, but this was based on the fact that small mlABPs can compute a polynomial that is full-rank with respect to a suitable restricted subfamily of all equipartitions, while small multilinear formulas cannot.

Known lower bounds on mlABP size. While Theorem 1 shows that full-rankness is not sufficient to prove a superpolynomial multilinear circuit lower bound, it could potentially prove such a lower bound on mlABP size. However, the best mlABP lower bound we have (which is also based on this method) is only $\tilde{\Omega}(n^2)$. We review this result [16, 2] below, as it naturally leads to our new approach to mlABP lower bounds.

Assume that a full-rank polynomial $P(x_1, \ldots, x_n)$ has an mlABP of size s and assume for simplicity that the mlABP is "layered" in the sense that vertices are divided into layers and edges go from one layer to the next, and for all vertices v at layer i, the set X_v of variables seen on the paths from a to v is a set of size i. To prove a lower bound for such an mlABP, we note that for any $i \in [n]$, the polynomial P can be decomposed as

$$P = \sum_{v \text{ in layer } i} L_v(X_v) \cdot R_v(X \setminus X_v), \tag{1}$$

[2] It should be noted that *separations* between different multilinear models require us to sometimes use specializations of this technique that yield a more fine-grained understanding. But the basic technique suffices to just prove a lower bound.

[3] Raz's work actually proves this result for a suitable transcendental extension of any field but this can be circumvented over an infinite field by replacing the transcendental elements by random elements from (a large subset of) $\mathbb{F}$. For example, as we did in the proof of Theorem 26.

where L_v corresponds to the sum of all paths from a to v and R_v the sum of all paths from v to b. Proving a lower bound on the number of terms in any such decomposition yields a lower bound on the number of vertices at layer i.

To lower bound the number of terms in the decomposition, we turn to *discrepancy*. We show that if the number of summands on the right hand side of (1) is small, then we can find an equipartition $\Pi = (Y, Z)$ such that each summand has low-rank for this choice of Π. This can be guaranteed by ensuring that, for each v, the induced partition $\Pi' = (Y \cap X_v, Z \cap X_v)$ of X_v is sufficiently "imbalanced". Informally, this ensures that the coefficient matrices of L_v and R_v are both "far from square" matrices and hence that $M_\Pi(f_v \cdot g_v)$ is far from full-rank.

This leads to the combinatorial question of understanding the smallest set system where we have such a discrepancy bound [16]. More precisely, fix an $i \in [n]$. Call a set system $\mathcal{Y} \subseteq \binom{[n]}{i}$ *k-balancing* if for any balanced function $f : [n] \to \{\pm 1\}$ (representing the equipartition), there is a $Y \in \mathcal{Y}$ such that $|\sum_{y \in Y} f(y)| \leq k$. Then, the size of the smallest such k-balancing set system represents the best lower bound we can prove for the number of vertices in layer i of the mlABP. Note that the answer to this is at most n, since the set of all intervals in $\{1, \ldots, n\}$ of size i yields a set system with discrepancy bound 1. This simple example is nearly tight for $i \in [n/4, 3n/4]$ and small k: Alon, Kumar and Volk [2] proved a tight lower bound of $\Omega(n/k)$ on the size of such "balancing set systems" for all k, which leads to a near-linear bound on the number of vertices in each layer and hence a near-quadratic lower bound on mlABP size.[4]

Our new technique. In order to prove a stronger lower bound, we suggest a modification of the decomposition in (1). We observe that such a decomposition of P can be obtained by looking at *any* vertex cut in the mlABP separating the source a from the sink b. As long as we can find a partition Π and a suitable cut C such that for each vertex $v \in C$, the set X_v is imbalanced with respect to Π, we obtain a lower bound. This gives us potentially exponentially many possibilities to work with and hence a much better chance at proving a lower bound against mlABPs.

By graph-theoretic arguments, the non-existence of such a cut is equivalent to the existence of an a-b path such that all the vertices v along the path are balanced. These vertices correspond to a *chain* of subsets of the variable set, which naturally brings us the following definition.

▶ **Definition 2** (Balanced-chain set systems). *Let X be a finite set with n elements and let $\mathcal{X} \subseteq \mathcal{P}(X)$ be a family of subsets of X. Denote by $\mathcal{C}(\mathcal{X})$ the set of* maximal chains *of $\mathcal{P}(X)$ contained in $\mathcal{X}$; i.e., a chain $(C_0, \ldots, C_l)$ is in $\mathcal{C}(\mathcal{X})$ iff $l = n$ and $C_i \in \mathcal{X}$ and $|C_i| = i$ for each $i \in \{0, \ldots, l\}$.*

For a function $f \colon X \to \{-1, 1\}$ which we will call a partition*, and $S \subseteq X$, let $f(S) := \sum_{x \in S} f(x)$. We say that f is* balanced *(or that f is a* balanced partition *of X) if $f(X) = 0$. We define the* chain-balance *of $\mathcal{X}$ with respect to f as*

$$\mathrm{cbal}_{\mathcal{X}}(f) := \min_{(C_0, \ldots, C_n) \in \mathcal{C}(\mathcal{X})} \max_{i \in [n]} |f(C_i)|.$$

We define the chain-balance of $\mathcal{X}$ to be

$$\mathrm{cbal}(\mathcal{X}) := \max_{\text{balanced } f \colon X \to \{-1, 1\}} \mathrm{cbal}_{\mathcal{X}}(f).$$

We say that $\mathcal{X}$ is a k-balanced-chain set system *if* $\mathrm{cbal}(\mathcal{X}) \leq k$.

[4] The general "unlayered" case is more complicated.

An example: intervals. Let $X := [n]$ and consider the set system $\mathcal{I}$ of all intervals contained in $[n]$:

$$\mathcal{I} := \{[i,j] \mid 1 \le i,j \le n\} \subseteq \mathcal{P}(X).$$

As noted above, this set system provides minimum-sized examples of 1-balanced set systems contained in $\binom{[n]}{i}$ for any i. It is therefore natural to understand if they yield balanced-*chain* set systems (the key new word is "chain"). The following balanced partition is an example of a partition such that $\mathcal{I}$ has large chain-balance: let $f : [n] \to \{\pm 1\}$ given by

$$f(j) := \begin{cases} 1 & \text{if } j \in [n/4] \cup [(3n/4)+1, n], \\ -1 & \text{otherwise.} \end{cases}$$

Pictorially, the graph of $i \mapsto f([i])$ comprises a mountain of linear height and a valley of linear depth. It can be shown that any chain in $\mathcal{C}(\mathcal{I})$ contains a set S that is very much not balanced; i.e., $|f(S)| \ge \Omega(n)$. Note that balance $n/2$ is the worst possible. This partition shows that the intervals set system $\mathcal{I}$ is not a balanced-chain set system, but it could be the case that this partition is an untypical example and $\mathcal{I}$ has small balance for most balanced partition. Our Theorem 8 shows that this is not the case: for most balanced partitions, $\mathcal{I}$ has chain-balance at least $n^{\Omega(1)}$.

1.1 Our results

Balanced-chain set systems and lower bounds. Our first result relates balanced-chain set systems to the problem of proving mlABP lower bounds via the min-partition rank method. We show the following two-directional equivalence.

▶ **Theorem 3.** *Let $s := s(n)$ be any growing function of n with $s \ge n$. Then the following hold:*

- *If any $(\log s)$-balanced-chain set system has size at least s, then any mlABP computing a full-rank polynomial has size at least $s^{\Omega(1)}$.*
- *Conversely, if there is an $O(1)$-balanced-chain set system of size s, then there is an mlABP computing a full-rank polynomial of size at most $s \cdot \mathrm{poly}(n)$.*

The reader may find the forward direction fairly intuitive, given the discussion above. In the other direction, it is possible to turn this purely combinatorial construction back into a polynomial that satisfies the algebraic property of full-rankness.

▶ **Observation 4.** *The above theorem provides a near-tight characterization of the size of the smallest mlABP in terms of the size of the smallest balanced-chain set system. The two directions differ only in the balance parameter: e.g., an mlABP lower bound of the form $s \ge n^c$ requires us to show that every $\Omega(c \log n)$-balanced-chain set system needs size n^c, but an upper bound of $s \le n^c$ can only be ensured by an $O(c)$-balanced-chain set system of size n^c. We can close this gap completely for the more restrictive setting of* set-multilinear *ABPs (defined below) where both directions use a balance parameter that is $\Theta(c)$. So at least in the set-multilinear setting, where the lower bound question is still open [4], we have a perfect characterization of the ABP complexity of full-rank polynomials in terms of the combinatorial question.*

The next question is to understand how to analyze and prove lower and upper bounds on the minimal size of a balanced-chain set system. A priori, this problem could be difficult, since finding a specific partition $f : X \to \{\pm 1\}$ with respect to which $\mathrm{cbal}_{\mathcal{X}}(f)$ is large could be

hard. However, we show, via a simple covering argument, that a *uniformly random* partition is as hard as a worst-case partition in the following sense: if there is a small set system that contains a balanced chain for most partitions, then there is also a small balanced-chain set system. We make this precise below.

▶ **Definition 5.** *For $\varepsilon \in [0,1]$ and $k \in \mathbb{R}_{\geq 0}$, we call an mlABP set system $\mathcal{X}$ an (ε, k)-*balanced-chain set system *if, for a uniformly random balanced partition $f : X \to \{\pm 1\}$, we have*

$$\mathbb{P}_f \left[\mathrm{cbal}_{\mathcal{X}}(f) \leq k\right] \geq \varepsilon.$$

Note that a k-balanced-chain set system is also a $(1,k)$-balanced-chain set system.

▶ **Lemma 6** (Worst-case to average-case reduction)**.** *If there is an (ε, k)-balanced-chain set system, over the set $[n]$, of size s, then there is a k-balanced-chain set system, over the set $[n]$, of size at most $O(sn/\varepsilon)$.*

We use both directions of Theorem 3 and Lemma 6 in what follows.

An improved upper bound on the size of balanced-chain set systems. We note that Theorem 3 in combination with the results of Raz (Theorem 1) immediately implies that there is an $O((\log n)^2)$-balanced-chain set system of size $n^{O(\log n)}$. In fact, it is not hard to give a direct proof of the slightly stronger fact that there is a 1-balanced-chain set system of size $n^{O(\log n)}$. This can be proved using an inductive construction (for completeness, see Theorem 12), using the fact (already mentioned above) that the family of intervals contained in $[n]$ are tight for the "balancing set systems" problem. The above lemma, however, is still consistent with the fact that the mlABP complexity of any full-rank polynomial could be $n^{\Omega(\log n)}$, which is no better than the upper bound in Theorem 1. Using a more sophisticated inductive argument, we are able to improve the easy upper bound superpolynomially, marking the first such improvement.

▶ **Theorem 7.** *There is a 1-balanced-chain set system of size at most $n^{O(\log n/\log\log n)}$.*

Using Theorem 3, this implies that there are full-rank mlABPs of size $n^{O(\log n/\log\log n)}$. An immediate consequence (using Theorem 1) is that mlABPs are superpolynomially more powerful than multilinear formulas.

This recovers a superpolynomial separation between these models that was proved in [7]. Our result is weaker than this work both quantitatively (the superpolynomial lower bound) and qualitatively (the construction of the set system and hence the mlABP are not uniform). However, we note that our proof is very different from that of [7]. While the proof of [7] reveals more about the structure of multilinear formulas by showing that they are not full-rank with respect to a much more restricted family of equipartitions (called arc-partitions), our proof yields a better understanding of the computational power of mlABPs, which is what we need to prove lower bounds against the latter model.

From a broader computational complexity perspective, understanding the size of ABPs is closely related to studying the computational power of dynamic programming. The construction above shows that dynamic programs can be quite powerful. More concretely, consider the following non-deterministic read-once branching program model. The devices in this model are directed acyclic graphs with a designated source a and a designated sink b. The input is $x \in \{0,1\}^n$ for even n. Each edge in the graph is labeled by two variables (x_i, x_j). On every directed path every variable appears at most once. Assume that all a-to-b

paths are of length exactly $n/2$. On input $x \in \{0,1\}^n$, an edge labeled by (x_i, x_j) is open iff $x_i + x_j = 1$ (in other words, it is balanced). The program accepts x iff there is an open path from a to b. We care about such programs that accept x iff the Hamming weight of x is exactly $n/2$. A program that achieves this functionality can actually "prove" its correctness. If it accepts an input x, then the open path that verifies acceptance shows that half of the entries of x are 1s and half are 0s.

What is the smallest program with this functionality? It is not hard to inductively construct such a program of size $n^{O(\log n)}$. The above theorem gives a non-trivial *superpolynomial* improvement. The exact complexity remains an open question.

Intervals and random partitions. As outlined above, set systems based on intervals have played an important role in previous investigations on multilinear lower bounds. Such families are optimal for constructions of balanced set systems and furthermore, results that separate multilinear circuits and ABPs from formulas [14, 7] are also based on combinatorial constructions that use intervals.

It is therefore natural to ask if intervals can be used to construct k-balanced-chain set systems for small k. We have already seen before that this family does not yield a good balanced-chain set system for a worst-case f. However, Lemma 6 shows that if this set system yields a balanced chain for a random partition, then that would allow us also to devise a somewhat larger set system that has balanced chains for all partitions. We are able to rule out this possibility with a strong negative result. This provides evidence that our technique may be able to prove superpolynomial lower bounds against general mlABPs.

▶ **Theorem 8** (Intervals vs. random partitions). *The intervals set system $\mathcal{I}$ defined before is not an (ε, k)-balanced-chain set system for any $\varepsilon > 2^{-\Omega(n^{1/5})}$ and $k < n^{1/5}$.*

In particular, the above is a considerable generalization of a recent result of Chatterjee, Kush, Saraf and Shpilka [4], whose results imply such a bound for the subset $\mathcal{I}_0$ of $\mathcal{I}$ that consists only of intervals of the form $[i]$ for $i \in \{0, \ldots, n\}$. The set system $\mathcal{I}_0$ contains just one maximal chain, while the set system $\mathcal{I}$ contains *exponentially* many chains, which makes the arguments from [4] inapplicable for $\mathcal{I}$. Using our results, we can prove the following lower bound for a generalization (formally defined in Section 5.4) of the sum of ordered set-multilinear ABPs model considered by [4].

▶ **Corollary 9** (Lower bounds for sum of interval-mlABPs). *Let $\mathbb{F}$ be any field. Any $\Sigma_\pi$mlABP program computing a full-rank polynomial $P(x_1, \ldots, x_n) \in \mathbb{F}[x_1, \ldots, x_n]$ must have size at least $2^{\Omega(n^{1/5})}$.*

The main technical part of the proof of Theorem 8 analyzes the discrete Fréchet distance between two one-dimensional random walks. This is a basic problem of independent interest. We have two random walks, and we want to understand what is the best time parametrization that makes them as close as possible. Here is a slightly more formal description (but still not fully formal; for details see Section 4). A random walk defines a random function $f : \{0, 1, \ldots, n\} \to \mathbb{Z}$ where $f(t)$ is the position of the walk at time t. The Fréchet distance between two functions f and g is

$$d_F(f, g) := \min_{\alpha,\beta} \max_t |f(\alpha(t)) - g(\beta(t))|,$$

where α, β are time parameterizations (that may depend on both f and g). We have two independent (1-Lipschitz) random functions, we want to prove that we cannot hope to parametrize time to make them close (even when we know the future). The property that makes this problem challenging is the dependency on the future.

The set-multilinear setting. Two recent works have conducted investigations into the power of *set-multilinear* ABPs (smlABPs). It is not hard to show that proving lower bounds against mlABPs is at least as hard as the analogous problem for smlABPs, and proving lower bounds against the latter is still an open question. The work of Bhargava, Dwivedi and Saxena [3] showed that proving lower bounds on smlABPs for low-degree polynomials implies general ABP lower bounds. A follow-up result of [4] showed restricted lower bounds against smlABPs in both the low and high-degree settings. As mentioned above, this latter paper implicitly proves a result that is a special case of Theorem 8 above.

1.2 Outline of the proofs

From ABPs to balanced-chain set systems and vice versa. We sketch the forward direction of Theorem 3 by showing if there is an mlABP of size s computing a full-rank polynomial P, then there is an $O(\log s)$-balanced-chain set system of size at most s. Let X denote the set of n variables. We associate with each vertex v in the mlABP a set $X_v \subseteq X$ of variables that appear on paths from the source vertex to v. The set system $\mathcal{X}$ we consider is the set of all such X_v (there are some hidden technicalities here). If there is a balanced $f : X \to \{\pm 1\}$ such that there is no $O(\log s)$-balanced-chain in $\mathcal{X}$, then we can find a cut C separating source from sink in the mlABP containing only vertices v such that X_v is, for example, $10 \log s$-imbalanced with respect to f. Using the decomposition

$$P = \sum_{v \in C} L_v(X_v) \cdot R_v(X \setminus X_v),$$

we can show that P is not full-rank with respect to f, which is a contradiction.

In the setting of *set-multilinear* ABPs, we can get a tighter result by showing that there is an $O(\log_m s)$-balanced-chain set system of size at most s where m is the number of variables in each part of the set-multilinear variable-set.

For the other direction, we start with a k-balanced-chain set system $\mathcal{X}$ of size s and show that there is a combinatorial transformation that produces a 1-balanced-chain set system of size $s \cdot n^k$. In the 1-balanced-chain setting, we use an idea that was first used by Raz [14] to construct full-rank polynomials. The observation is that, for a known equipartition $\Pi = (\{y_1, \ldots, y_{n/2}\}, \{z_1, \ldots, z_{n/2}\})$ of the n variables, the simple "gadget" polynomial

$$Q_\Pi = \prod_{i=1}^{n/2} (y_i + z_i)$$

is full-rank with respect to Π. To handle an unknown partition, we construct a polynomial that is a sum of such polynomials for all possible equipartitions. A 1-balanced-chain set system gives us a recipe for constructing such a polynomial efficiently. Given any equipartition $\Pi = (Y, Z)$, we know that there is a chain $C = (C_0, \ldots, C_n)$ of sets in the set system such that each set has chain-balance either 0 or 1 with respect to Π. In particular, the even-sized sets $C_0, C_2, C_4, \ldots, C_n$ all are perfectly balanced and $C_{i+2} = C_i \cup \{y_{i/2+1}, z_{i/2+1}\}$ for each even $i < n$. A natural choice to construct a full-rank Q_Π is to pair up $y_{i/2+1}$ with $z_{i/2+1}$. It is possible to design an mlABP that computes the sum of all such Q_Π for all chains in a given set system, as in [7]. To ensure that the resulting polynomial is full-rank with respect to all equipartitions, we need to take a suitable linear combination of all these Q_Π. Raz [14] observed that this can be done over a transcendental extension of the base field $\mathbb{F}$ by using these additional transcendental field elements in the linear combination (the above gadget polynomial is actually from [16]). We note that if the field $\mathbb{F}$ is infinite, these transcendental

elements can be replaced by random elements of $\mathbb{F}$, giving us a full-rank polynomial over the base field. A similar construction (with a different gadget polynomial) also works in the set-multilinear case.

Worst-case to average-case reduction. Assume that we have an (ε, k)-balanced-chain set system $\mathcal{X}$. The main observation is that, given a worst-case balanced partition $f : X \to \{\pm 1\}$, and a uniformly random permutation $\pi : X \to X$, the function $f \circ \pi$ is a uniformly random partition and hence, the set system $\mathcal{X}$ contains a k-balanced-chain with respect to $f \circ \pi$ with probability at least ε. This is the same as saying that the random set system $\mathcal{X}_\pi$ (obtained by permuting the elements of each set in $\mathcal{X}$ according to π) contains a k-balanced-chain with respect to f with probability at least ε. Choosing $O(n/\varepsilon)$ many such permutations π and taking the union of the corresponding set systems allows us to increase this probability to $1 - 2^{-n}$, at which point the probabilistic method guarantees the existence of a suitable set system that works for all partitions.

Constructing a 1-balanced-chain set system of size $n^{O(\log n/\log\log n)}$. By the preceding argument, it suffices to construct a 1-balanced-chain set system of this size that contains a 1-balanced-chain for a uniformly random balanced partition f with good probability (e.g. $1/\operatorname{poly}(n)$). To do this, we identify the set X with $[n]$ and view $f : X \to \{\pm 1\}$ as a uniformly random walk of length n on the integers that starts and ends at the origin 0 (also known as a random bridge). We note that a maximal chain that is balanced with respect to f contains $n/2$ sets C that are balanced with respect to f. A random walk that returns to 0 many times would give us many *intervals* that satisfy this criterion. A result of Csáki, Erdős and Révész [6] tells us that a uniformly random walk f, with probability $\geq 1/\operatorname{poly}(n)$, returns to 0 at least once every $m \approx n/\log n$ steps. This implies that the set of all intervals, while not balanced, does contain a *non-maximal* chain of balanced intervals

$$\varnothing \subseteq [i_1] \subseteq [i_2] \subseteq \cdots \subseteq [i_r] = [n]$$

where each subsequent interval in the chain adds at most m elements to the previous one. To finish the argument, we only need to add sets to "fill in" the missing sets in the chain to make it maximal. This can be formulated as the same problem for the smaller set size m. Implementing a recursive construction leads to a set system of the given size.

▶ **Observation 10.** *While we only used the family of intervals in the above construction, it should be noted that the overall family constructed above consists of sets that are very different from intervals. This is because we need to carry out the aforementioned average-case to worst-case argument at each stage, which is done by randomly permuting the elements in the sets. This does not preserve the property of being an interval. It seems interesting to derandomize this "construction".*

Intervals and random partitions. Let $\mathcal{I}$ denote the intervals set system over $[n]$ as defined before. Assume that $f : [n] \to \{\pm 1\}$ is chosen uniformly at random (for the sake of exposition, we drop for now the assumption that f is balanced). The claim is that the chance that there is a maximal chain $(C_0, \ldots, C_n)$ of intervals that make f balanced is exponentially small.

For the sake of exposition, we also assume that the chain starts at $C_1 = \{n/2\}$. The random walk f can be partitioned into two independent random walks: g that goes from $n/2$ to n and h that goes from $n/2$ to 1. The random walk f is chain-balanced only if there are time parameterizations under which g, h are close at all times (this is a discrete version of

the Fréchet distance). We thus have two independent random walks (each of length $n/2$) and the claim is that it is very unlikely that there is a way to parametrize time so the distance between g, h is say $k < n^{1/3}$. The main idea is to first sample the random walk g and identify a pattern in it that h can not possibly follow. Fix the position of the random walk g at times that are multiples of some ℓ. Let p_t be the position of g at time $t\ell$ for $t \leq T := n/(2\ell)$. Typically, the distance between p_t and p_{t+1} is roughly $\sqrt{\ell}$. For concreteness, think of the case that $p_t = 0$ for even t and $p_t = \sqrt{\ell}$ for odd t. It remains to show that there is no sequence of (strictly increasing) time steps $\tau_1 < \tau_2 < \ldots < \tau_T$ such that $h(\tau_t)$ is close to p_t for all t.

Let $(C_0, \ldots, C_n)$ denote any maximal chain in $\mathcal{I}$ and assume that $C_1 = \{i\}$. We can decompose f into two random functions $g : [i] \to \{\pm 1\}$ and $h : [i+1, \ldots, n] \to \{\pm 1\}$. Each set C_j in the chain can be written as the disjoint union of two intervals $I'_j = [i - i' + 1, i]$ and $I''_j = [i+1, i+j-i']$ contained in the domains of g and h respectively. The quantity $f(C_j)$ is $g(I'_j) + h(I''_j)$ and thus the balance of this chain with respect to f is equal to

$$k := \max_j |g(I'_j) + h(I''_j)| = \max_j |g(I'_j) - (-h(I''_j))|.$$

Thinking of g and $-h$ as random walks on the integers, the chain can be interpreted as an "alignment" of these two random walks that ensures that they stay within distance k at all time steps, or equivalently, these two random walks are at distance at most k from each other in *discrete Fréchet distance* (defined formally below). The main technical part of the proof is to show that the probability that $k \leq n^{O(1)}$ is exponentially small.

This is done by finding some pattern in one of the two walks, say g, that does not appear in the other except with small probability. To simplify further, assume that g and h are random walks of the same length $n/2$. We divide the time steps of g into epochs of some suitable length ℓ and consider the positions of the random walk g at these various time steps. I.e.

$$p_1 = g([(n/2) - \ell, n/2]), p_2 = g([(n/2) - 2\ell, n/2]), \ldots, p_{n/2\ell} = g([n/2]).$$

As long as ℓ is large enough, there is enough randomness between any pair of time steps to ensure that most pairs of consecutive positions are quite far from one another.[5] Conditioning on these positions, we show that with high probability, there is no subsequence of (strictly increasing) timesteps when the positions of $-h$ are $p'_1, p'_2, \ldots, p'_{n/2\ell}$ such that $|p_t - p'_t| \leq k$ for each t. This implies that the two walks cannot be aligned as described above.

To see that $-h$ cannot come close to the positions achieved by g, we think of the case when $p_1 = p_3 = p_5 = \cdots = 0$ and $p_2 = p_4 = p_6 = \cdots = \sqrt{\ell}$. Intuitively, this is a bad case as 0 is the most probable location of a random walk at any (even) time step and we expect a random walk to travel a distance of $\sqrt{\ell}$ every ℓ steps. Now, the expected number of times that a random walk of length $n/2$ returns to 0 is $\Theta(\sqrt{n})$ and we can also prove a concentration bound on the upper tail of this random variable. Thus, if we choose ℓ significantly smaller than $\sqrt{n}$, it is very unlikely that the random walk $-h$ can return to (a position close to) 0 at least $\Omega(n/\ell)$ many times.

1.3 Conclusion and further questions

In this work, we study the power and limitations of the min-partition rank method to prove lower bounds for mlABPs. We first introduce a combinatorial complexity measure, the chain-balance, for set systems that generalizes previous combinatorial notions used to

[5] In the actual argument, we vary the lengths of these epochs to ensure that *every* consecutive pair of positions are indeed far from one another.

prove lower bounds for mlABPs. We prove general properties of this measure, together with an $n^{O(\ln n/\ln\ln n)}$-upper bound for the size of a set system with chain-balance 1, and an average-case $\Omega(n^{1/5})$-lower bound for the chain-balance of restricted set systems obtained from intervals. Then we translate those combinatorial results into the algebraic complexity world. We obtain both upper and lower bounds for mlABP and its variants. We show that no full rank polynomial can be computed by a $\sum_\pi$ mlABP program of size $2^{o(n^{1/5})}$, and also recover a superpolynomial separation between multilinear formulas and mlABPs. Additionally, our $n^{O(\log n/\log\log n)}$ upper bound on the size of mlABP computing a full rank polynomial implies that the min-partition rank method cannot be used to prove the potentially optimal $n^{\Omega(\ln n)}$-size lower bound against mlABPs for a (full rank) polynomial in **mlVP**. The main open questions related to our work are the following:

- Can we close the gap between the lower bound $\Omega(n^2/k)$ and upper bound $n^{O(\ln n/\ln\ln n)}$ for the size of k-balanced-chain set systems? Both directions are interesting since Theorem 3 shows that sufficiently high lower bounds on the size of balanced-chain set systems imply new lower bounds for the size of mlABPs computing full rank polynomials, and a poly(n) upper bound implies that the min-partition rank method cannot be used to prove superpolynomial lower bounds for mlABPs.
- Can we prove $n^{\Omega(1)}$ lower bounds for the chain-balance of other families of restricted set systems? In Theorem 8, we showed an $\Omega(n^{1/5})$-lower bound for the chain-balance of intervals set systems, and applied it almost as a black box to obtain an exponential lower bound against the $\sum_\pi$ mlABP model. Thus, it is natural to ask whether we can reduce the problem of showing lower bounds for other interesting restricted versions of mlABPs to the question of showing lower bounds for the chain-balance of restricted set systems.
- Can we obtain an explicit construction of full rank mlABPs of size $n^{o(\ln n)}$? A possible approach for this question is a derandomization of our construction. In particular, if we are allowed to use field extensions (as the constructions in [15, 7]), it is sufficient to derandomize our proof of the worst-case to average-case reduction, which is a covering argument via random permutations of a given set system (Lemma 14).

1.4 Organization of the paper

We start with some preliminaries and then prove some combinatorial properties of the chain-balance of set systems in Section 2, including the worst-case to average-case reduction (Lemma 6). We will then show our main combinatorial theorems, starting with the construction of a 1-balanced-chain set system of size $n^{O(\log n/\log\log n)}$ (Theorem 7) in Section 3, followed by the analysis of intervals with respect to random partitions (Theorem 8) in Section 4. Finally, in Section 5, we prove the formal connection between the combinatorial constructions and the mlABP complexity of full-rank polynomials (Theorem 3) and derive the consequences of our combinatorial theorems to algebraic complexity, that is, a separation between multilinear formulas and mlABPs, and an exponential lower bound for the $\sum_\pi$ mlABP model (Corollary 9).

The omitted proofs can be found in the full version of this paper (`https://eccc.weizmann.ac.il/report/2026/001/`).

2 Properties of balanced-chain set systems and their chain-balance

In this section, we prove some properties of balanced-chain set systems related to bounds for the size of these systems and conversions between different notions of chain-balance.

2.1 Upper and lower bounds for the size of balanced-chain set systems

Now let us define a natural class of set systems.

▶ **Definition 11.** *For every $n, m \in \mathbb{N}$, we denote by $\mathcal{I}_{n,m}$ the* m-interval set system *over $[n]$ defined as*

$$\mathcal{I}_{n,m} := \{\, I_1 \cup \ldots \cup I_l \mid I_1, \ldots, I_l \text{ are intervals of } [n] \text{ and } l \leq m \}.$$

Assume that $m \leq n/2$. Since any element of $\mathcal{I}_{n,m}$ is specified by the endpoints of the intervals, using the usual bounds on the binomial numbers, we get that the size of the m-interval set system is at most

$$|\mathcal{I}_{n,m}| \leq (n/m)^{O(m)}.$$

We note below that using a simple inductive analysis, we can obtain a 1-balanced-chain set system from intervals.

▶ **Theorem 12.** *For every even number $n \in \mathbb{N}$, the chain-balance of the $(2\lceil \lg n \rceil)$-interval set system is* 1.

Proof. We omit this proof but it can be found in the full version of this paper. Actually, in Section 3, we analyse the probabilistic behaviour random balanced partitions, instead of the structural result used in the proof of this theorem, and construct a 1-balanced-chain set system of size $n^{O(\lg n / \lg \lg n)}$. ◀

▶ **Observation 13.** *We also note that using previous results about balancing systems, we can prove a $\Omega(n^2/k)$ lower bound for the size of any k-balanced-chain set system for $k \leq n/5$. More specifically, a result by Alon, Kumar, and Volk [2] can be rephrased as follows: for every $k \in \mathbb{N}$ and sufficiently large $n \in \mathbb{N}$, if a collection $S_1, \ldots, S_m \subseteq [n]$ of sets satisfies that $2k \leq |S_i| \leq n - 2k$ for every $i \in [m]$ and, for every balanced partition f of n, there is an $i \in [m]$ such that $|f(S_i)| \leq 2k$, then*

$$m \geq \Omega(n/k).$$

Hence, for any $2k$-balanced-chain set system $\mathcal{X}$ over $[n]$, we can consider, for every $l \in \{2k, \ldots, n-2k\}$, the collection

$$\mathcal{X}_l := \{\, S \in \mathcal{X} \mid |S| = l \}.$$

As $\mathrm{cbal}(\mathcal{X}) \leq 2k$, every balanced partition f has a maximal chain $(C_0, \ldots, C_n)$ of $\mathcal{X}$ such that, for every $i \in [n]$, the set C_i satisfies $|C_i| = i$ and

$$|f(C_i)| \leq 2k.$$

By the lower bound above, we obtain that $|\mathcal{X}_l| \geq \Omega(n/k)$, which implies that

$$|\mathcal{X}| \geq \Omega((n-4k)n/k).$$

2.2 Conversion of (p, l)-balanced-chain systems into l-balanced-chain systems

As mentioned in the introduction, we can prove a worst-case to average-case reduction for balanced-chain set systems.

▶ **Lemma 14** (Lemma 6). *Let $n \in \mathbb{N}$ be an even number, and X be a set of n elements. Let $l \in [n]$ and $p \in (0, 1]$. If $\mathcal{X}$ is a (p, l)-balanced-chain set system, over X, of size s, then there is an l-balanced-chain set system $\mathcal{Y}$, over X, of size $O(sn/p)$.*

2.3 Conversion of l-balancedness to 1-balancedness

Let us now prove a combinatorial lemma that shows how to convert an l-balanced-chain set system into a 1-balanced-chain set system with some size blow-up.

▶ **Lemma 15.** *Let $n \in \mathbb{N}$ be an even number, and let X be a finite set with n elements. If $\mathcal{X}$ is an l-balanced-chain set system, over X, with size s, then there is a set system $\mathcal{Y}$ over X such that*

- *$\mathcal{Y}$ has chain-balance 1; and*
- *the size of $\mathcal{Y}$ is at most $s\binom{|X|}{\leq l}$.*

3 Construction of a 1-balanced-chain set system of size $n^{O(\ln n/\ln\ln n)}$

In this section, our goal is to prove Theorem 7. Our proof uses concepts from one-dimensional random walk theory (Definition 16) and a result about the probability distribution of the length of the longest excursion in a random bridge (Lemma 17). In Section 5.2, the main result of this section is used to obtain a separation between multilinear formulas and mlABPs.

▶ **Definition 16.** *Let $n \in \mathbb{N}$ and let $f\colon [n] \to \{-1,1\}$. We associate a function $W_f\colon \{0,\dots,n\} \to \mathbb{Z}$ to f as follows: let $W_f(0) := 0$ and, for every $i \in [n]$,*

$$W_f(i) := f([i]),$$

where $f([i]) = \sum_{j\in[i]} f(j)$ as defined in Definition 2. We call W_f the walk defined from f, *and n the* length *of W_f. Note that the definition above is a bijection between f and W_f. Thus, for a given walk W, we denote by f_W the function such that $W = W_{f_W}$. Sometimes we also consider W_f to be the sequence $(W_f(0),\dots,W_f(n))$.*

We say that W_f is a bridge *if f satisfies $f([n]) = 0$. We say that W_f is a* excursion *if W_f is a bridge and, for every $i \in [n-1]$,*

$$W_f(i) \neq 0.$$

For f sampled uniformly at random from the set of all functions from $[n]$ to $\{-1,1\}$, we say $W := W_f$ is a random walk *of length n, that is,*

$$W \sim \mathrm{Unif}(\{\, W_f \mid f\colon [n] \to \{-1,1\}\}.$$

Simarly, for even n, we define a random bridge *B of length n as*

$$B \sim \mathrm{Unif}(\{\, W_f \mid f\colon [n] \to \{-1,1\}, f([n]) = 0\}).$$

Let B be a random bridge of length n and $f := f_B$. We say that a point $t \in \{0,\dots,n\}$ is a zero *of B if*

$$f([t]) = 0,$$

or, equivalently, $B(t) = 0$. Let $Z(B)$ be the (random) set of zeros of B. Note that $\{0,n\} \subseteq Z(B)$ by the definition of bridge. Let $\{z_0,\dots,z_l\} := Z(B)$ be the elements of $Z(B)$, where $z_0 < z_1 < \ldots < z_l$. Let $\lambda(B)$ be the random variable defined as follows:

$$\lambda(B) := \max_{i\in[l]} z_i - z_{i-1}.$$

We call $\lambda(B)$ the length of the longest excursion *in B.*

▶ **Lemma 17** (Lemma 10 from [6]). *Let $n \in \mathbb{N}$ and let B be a random bridge of length $2n$. There are universal real constants $0 < C_1 \leq C_2$ and $\beta > 0$ such that, for every $a \geq n^{2/3}$, we have*

$$\mathbb{P}_B[\lambda(B) \leq 2a] = C(n,a) \min\{(n+1)^{-1/2}, a^{-1/2}\} \exp(-\beta n/a),$$

for $C_1 \leq C(n,a) \leq C_2$.

Let us now prove our main theorem.

▶ **Theorem 18** (Theorem 7). *For every sufficiently large even number $n \in \mathbb{N}$, there is a 1-balanced-chain set system $\mathcal{X}_n$ over $[n]$ of size $n^{O(\ln n / \ln \ln n)}$.*

Proof. For every $n \in \mathbb{N}$, let $N(n)$ be the minimum size of a 1-balanced-chain set system over $[n]$, and let A_n be a fixed 1-balanced-chain set system of size $N(n)$. Our goal is to obtain a recursive upper bound for $N(n)$.

For any $m \in [n]$, let B_m be the set system defined as follows:

$$B_m := \{\, [1,i] \odot A_t \mid t \in [m], 0 \leq i \leq n-t \},$$

where we define $[1,i] \odot A_t$ as

$$[1,i] \odot A_t := \{\, [1,i] \cup \{\, i+j \mid j \in R\} \mid R \in A_t \},$$

which can be interpreted as the translation of the set system A_t to the end of the interval $[1,i]$. Note that B_m is a set system over $[n]$ of size at most $n^2 N(m)$. Let f be a uniformly random balanced partition of $[n]$. We say that a point $t \in [n]$ is a *zero* of f if $\sum_{i \in [t]} f(i) = 0$, and we say that a set system A is *f-balanced* if $\mathrm{cbal}_A(f) \leq 1$. Let E_m be the event that the maximum gap between two consecutive zeros of f is at most m, and let

$$\delta_m := \mathbb{P}_f[E_m].$$

Consider the following auxiliary lemma:

▶ **Lemma 19.** *For any balanced partition f of $[n]$, if $f \in E_m$, then B_m is f-balanced.*

Proof. This follows from the fact that if $f \in E_m$, then the maximum gap between two consecutive zeros z_i and z_{i+1} of f is at most m. Consequently, we can construct a f-balanced chain of B_m in the following inductive way:

1. We know that $z_1 \leq m$, and A_{z_1} has a $f\restriction_{[1,z_1]}$-balanced chain C_1, and $A_{z_1} = [1,0] \odot A_{z_1}$, so C_1 is also a subchain of B_m.
2. For every $i \geq 1$, suppose that we already have a subchain C_i of B_m such that C_i is $f\restriction_{[1,z_i]}$-balanced. Note that $f\restriction_{[z_i+1,z_{i+1}]}$ is a balanced partition of $[z_i+1, z_{i+1}]$, and, by defining a balanced partition g of $[z_{i+1} - z_i]$ as

   $$g(j) := f(z_i + j)$$

   for every $j \in [z_{i+1} - z_i]$, we can obtain a g-balanced chain C from the set system $A_{z_{i+1}-z_i}$. Furthermore, we have that $z_i + S \subseteq [z_i+1, z_{i+1}]$ for every $S \in C$, which implies that

   $$C_{i+1} := C_i \cdot ([1, z_i] \cup (z_i + S) \mid S \in C)$$

   is a $f\restriction_{[1,z_{i+1}]}$-balanced subchain of B_m, where $\cdot$ denotes the concatenation of chains (which can be interpreted as strings). By induction on i, we can obtain f-balanced chain of B_m. ◀

By Lemma 17 and interpreting its result about random bridges B in the language of random balanced partitions f, there are positive universal constants $b, c \in \mathbb{R}$ such that, for every even $m \in [n]$ such that $(n/2)^{2/3} \leq m/2$, we have

$$\delta_m = \mathbb{P}_B[\lambda(B) \leq m] \geq cn^{-1/2} \exp(-bn/m).$$

Now fix the value of m to $m := 2\lceil n/(2 \ln n)\rceil$, which implies that

$$\delta := \delta_m \geq c/n^{b+1/2}.$$

Hence, for this choice of m, we have that B_m is a $(\delta, 1)$-balanced-chain set system over $[n]$ by Lemma 19. By Lemma 14, there is a 1-balanced-chain set system A, over $[n]$, of size

$$|A| \leq O(|B_m|n/\delta) \leq (n^2 N(n/\ln n))n^{b+2} \leq n^d N(n/\ln n),$$

for a positive constant d. By solving the recurrence relation $N(n) \leq n^d N(n/\ln n)$, we obtain that

$$N(n) \leq n^{O(\ln n/\ln\ln n)}.$$ ◀

4 The chain-balance of interval set systems and the Fréchet distance of random walks

In this section, our main goal is to prove Theorem 8. In order to prove this theorem, we establish a connection between the chain-balance of interval set systems and the Fréchet distance between two random walks (Lemma 21), and prove that, with very high probability, the Fréchet distance of two random walks of length n is at least $n^{\Omega(1)}$ (Lemma 22). In Section 5.4, the main result of this section is used to obtain lower bounds on the size of sum of interval-mlABPs computing full rank polynomials.

4.1 Fréchet distance of random walks

Let us start by defining the Fréchet distance of two discrete functions [8].

▶ **Definition 20.** *Let $n \in \mathbb{N}$, and let $l, r \in [n]$ such that $l + r = n$. Let $X\colon \{0, \ldots, l\} \to \mathbb{R}$ and $Y\colon \{0, \ldots, r\} \to \mathbb{R}$ be two functions. We define the* (discrete) Fréchet distance $d_F(X, Y)$ *between X and Y as*

$$d_F(X, Y) := \min_{\alpha,\beta} \max_{t\in[n]} |X(\alpha(t)) - Y(\beta(t))|,$$

where the minimum ranges over all non-decreasing functions

$$\alpha\colon \{0, \ldots, n\} \to \{0, \ldots, l\} \text{ and } \beta\colon \{0, \ldots, n\} \to \{0, \ldots, r\}$$

satisfying the following properties:

- $\alpha(0) = 0 = \beta(0)$;
- $\alpha(n) = l$ *and* $\beta(n) = r$;
- *For every $t \in [n]$, we have $\alpha(t) \leq \alpha(t-1) + 1$, and $\beta(t) \leq \beta(t-1) + 1$, and*

$$\alpha(t) - \alpha(t-1) + \beta(t) - \beta(t-1) = 1.$$

Now let us state the main technical lemmas of this section.

▶ **Lemma 21** (Discrepancy and Fréchet distance). *Let $n \in \mathbb{N}$ be an even number, and let $\mathcal{I} := \mathcal{I}_{n,1}$ be the 1-interval set system (Definition 11). Let $l, r \in [n]$ such that $l + r = n$, and let $W_{l,r} := (W_l, W_r)$ be a pair of two independent random walks W_l and W_r (Definition 16) of length l and r, respectively. For f being a uniformly random balanced partition, and $W := (X, Y) := W_{n,n}$, and $X_l := X\restriction_{\{0,\ldots,l\}}$ and $Y_r := Y\restriction_{\{0,\ldots,r\}}$, we get that, for every $k \in \mathbb{R}$,*

$$\mathbb{P}_f[\mathrm{cbal}_{\mathcal{I}}(f) \leq k] \leq O(n^{5/2}) \max_{r\in[n]} \mathbb{P}_W[d_F(X_r, Y_{n-r}) \leq k].$$

▶ **Lemma 22** (Fréchet distance of two random walks). *There is a positive constant $c_1 \in \mathbb{R}$ such that, for every sufficiently large $n \in \mathbb{N}$, and every $l := l(n) \in [n]$, and every $\varepsilon := \varepsilon(n) \in (0, 1/4)$ such that $n^{1/4-\varepsilon} \geq \frac{2}{3} \ln n$, we have, for $(X, Y) := W := W_{l,n-l}$,*

$$\mathbb{P}_W[d_F(X, Y) < n^{1/4-\varepsilon}] \leq \exp(-c_1 \cdot \min\{n^{1/4-\varepsilon}, n^{4\varepsilon}\}).$$

Proof. This lemma is proved in Section 4.2. ◀

Let us now prove Theorem 8.

▶ **Theorem 23** (Theorem 8). *There is a universal constant $c > 0$ such that, for every sufficiently large even number $n \in \mathbb{N}$, the 1-interval set system $\mathcal{I}_{n,1}$ is not an (ε, k)-balanced-chain set system for $\varepsilon > 2^{-cn^{1/5}}$ and $k < n^{1/5}$.*

4.2 Proof of Lemma 22

We omit the proof of this lemma, but we give an informal description of its proof. We consider two random walks g and h of total length n. Assume w.l.o.g. that g has length at least $n/2$. Let Δ be a parameter that we will choose later. We observe that w.h.p. there are approximately $k \approx n/\Delta^3$ points in g such that every consecutive pair of points is Δ-far from each other. We then analyze the probability that h can reach close to each of these points (which is required for the Fréchet distance between them to be small). The number of steps to go from one of these points to the next is captured by a first passage random variable F_δ with distance parameter $\delta = \Theta(\Delta)$. For the Fréchet distance to be small, the sum of k independent copies of these random variables must be at most n. However, this probability is exponentially small as long as $k^2\delta^2 \gg n$ which is true for small enough Δ.

5 Full rank multilinear ABPs and balanced-chain set systems

In this section, our goals are to show a relationship between constructions of mlABPs computing full rank polynomials and balanced-chain set systems, and use this relationship together with our results from Sections 3 and 4 to, respectively, obtain a separation between the power of multilinear formulas and mlABPs to compute full rank polynomials, and prove lower bounds for the size of sums of *interval-mlABPs* computing full rank polynomials. We first prove our conversion from balanced-chain set systems to full-rank mlABPs (Section 5.1) and its consequences (Section 5.2), and then our conversion from full-rank mlABPs to balanced-chain set systems (Section 5.3) and its consequences (Section 5.4).

5.1 Construction of full rank mlABPs from balanced-chain set systems

In this section, our goals are to prove our main construction of full rank mlABPs from 1-balanced-chain set systems (Theorem 24) and generalize it to the case of l-balanced-chain set systems (Theorem 25). In order to obtain our generalization, we use a combinatorial lemma (Lemma 15) that shows how to convert an l-balanced-chain set system into a 1-balanced-chain set system with some size blow-up.

We start with the 1-balanced case. As is standard [14], we will start by proving it by enlarging the field by adding transcendental elements. Later on, we will show how to remove the transcendental elements at the expense of making the polynomial less "explicit" (by replacing these elements by random elements from the field).

▶ **Theorem 24.** *Let $n \in \mathbb{N}$ be an even number, and let X be a set of n elements. Let $W_X := \{ w_t \mid t := (i,u) \in [n] \times X \}$ and $V_X := \{ x_i \mid i \in X \}$ be two disjoint sets of variables with, respectively, n^2 and n elements. If $\mathcal{X}$ is a 1-balanced-chain set system over X, then there is a polynomial $P_{\mathcal{X}}$ over the variables $V_X \cup W_X$ such that the following hold:*

- *All coefficients of $P_{\mathcal{X}}$ are 0 or 1;*
- *For every field $\mathbb{F}$, the polynomial $P_{\mathcal{X}}$ is full rank over $\mathbb{F}(W_X)[V_X]$ and can be computed by an mlABP of size at most $|\mathcal{X}|$ over $\mathbb{F}(W_X)$.*

More generally, we have the following stronger version of Theorem 24 for l-balanced-chain set systems.

▶ **Theorem 25.** *Let $n \in \mathbb{N}$ be an even number, let $l \in [n]$ and X finite with n elements. Let $W_X := \{ w_t \mid t := (i,u) \in [n] \times X \}$ and $V_X := \{ x_i \mid i \in X \}$ be two disjoint sets of variables with, respectively, n^2 and n elements. If $\mathcal{X}$ is a l-balanced-chain set system over X, then there is a polynomial $P_{\mathcal{X}}$ over the variables $V_X \cup W_X$ such that*

- *All coefficients of $P_{\mathcal{X}}$ are 0 or 1;*
- *For every field $\mathbb{F}$, the polynomial $P_{\mathcal{X}}$ is full rank over $\mathbb{F}(W_X)[V_X]$ and can be computed by an mlABP of size at most $|\mathcal{X}|\binom{n}{\leq l}$ over $\mathbb{F}(W_X)$.*

5.2 A separation between multilinear formulas and mlABPs

As mentioned in the introduction, we can use Theorem 3 together with Theorem 18 to obtain a superpolynomial separation between multilinear formulas and mlABPs. We show this separation (Corollary 27) in this section, after we prove the following formal version of the second part of Theorem 3.

▶ **Theorem 26** (Second part of Theorem 3). *Let $n \in \mathbb{N}$ be an even number, let $l \in [n]$ and X finite with n elements. If $\mathcal{X}$ is a l-balanced-chain set system over X, then, for every infinite field $\mathbb{F}$, there is an n-variate full rank polynomial $P_{\mathcal{X}}$ over $\mathbb{F}$ that can be computed by an mlABP over $\mathbb{F}$ with size at most $|\mathcal{X}|\binom{n}{\leq l}$.*

Now let us use Theorem 26 and Theorem 18 to prove a separation between multilinear formulas and mlABPs.

▶ **Corollary 27.** *Let $n \in \mathbb{N}$ be a sufficiently large even number, and $\mathbb{F}$ be an infinite field. Then there is a n-variate polynomial P_n over $\mathbb{F}$ computed by a mlABP of size $n^{O(\ln n / \ln \ln n)}$, but any multilinear formula computing P_n has size at least $n^{\Omega(\ln n)}$.*

5.3 Construction of balanced-chain set systems from full-rank mlABPs

In this section, our goal is to prove our formal version of the first part of Theorem 3. In particular, we are going to prove a more general theorem (Theorem 28), which allows to relax the hypothesis from the first part of Theorem 3 to average-case almost-full-rank guarantees.

▶ **Theorem 28** (Balanced-chain set systems from layered mlABPs (first part of Theorem 3))**.** *Let $n \in \mathbb{N}$ be an even number, and X be a set of n variables. Let $p := p(n) \in [0,1], \varepsilon := \varepsilon(n) \in (0,1]$. Fix any mlABP $\mathcal{A}$ of size $s \geq n$ computing a (p,ε)-almost full-rank polynomial P over the variables X, Then there exists a $(p, O(\log(sn/\varepsilon))$-balanced-chain set-system $\mathcal{X}$ of size at most $O(s^2 n)$.*

5.4 Lower bounds against $\sum_\pi$ mlABPs

We show in this section how to use the combinatorial results in Section 4 and the connection between multilinear ABPs and balanced-chain set systems to obtain a stronger version of a result of [4]. While we state the result for multilinear ABPs, the same idea also works in the set-multilinear setting.

Let $X := \{x_1, \ldots, x_n\}$ be a set of n variables. Given a bijection $\pi : [n] \to [n]$, we denote by $\mathcal{I}_{X,\pi}$ *the set system consisting of intervals with respect to* π contained in the set X. Formally, we have

$$\mathcal{I}_{X,\pi} := \{\{x_{\pi(i)}, x_{\pi(i+1)}, \ldots, x_{\pi(j)}\} \mid 1 \leq i, j \leq n\}.$$

A π-interval mlABP is defined to be an ABP whose vertices are divided into $n+1$ layers. Each vertex v is labelled with a *π-interval* $Y_v \in \mathcal{I}_{X,\pi}$, and all vertices from the i-th layer are labelled with sets of size i. All edges go from vertices in layer i, for some i, to vertices in layer $i+1$. An edge (u,v) from vertex u in layer i to vertex v in layer $i+1$ is labelled by either a linear polynomial in the unique variable $x \in Y_v \setminus Y_u$, or a field constant. We say that an ABP $\mathcal{A}$ is an *interval mlABP* if there is a bijection π such that $\mathcal{A}$ is a π-interval mlABP.

We denote by $\Sigma_\pi$mlABP the computational model that is a sum of interval mlABPs (possibly with respect to different orderings π).[6] The size of an instance is the sum of all the sizes of the individual mlABPs that make up the instance.

This model is a generalization of the model considered in [4] who prove a lower bound for the sum of *ordered* (set-)multilinear ABPs. The difference between the ordered and interval settings is that in the former, all the X_v are intervals of the form $[i]$ for some $i \in \{0, \ldots, n\}$. Note that the underlying set system has only one maximal chain in the ordered case, but exponentially many chains in the interval setting. This makes our result below more challenging.

We show the following strong lower bound against $\Sigma_\pi$mlABP.

▶ **Theorem 29** (Corollary 9)**.** *Let $n \in \mathbb{N}$ be sufficiently large even number, and X be a set of n variables, and $\mathbb{F}$ be any field. Let P be a full rank polynomial in $\mathbb{F}[X]$. If $\mathcal{A}$ is a Σ_π*mlABP *of size s computing P, then*

$$s \geq 2^{\Omega(n^{1/5})}.$$

[6] The notation here is from [4] but we add an additional subscript "π" to recall that the mlABPs are all interval mlABPs.

5.5 The set-multilinear case

Both parts of Theorem 3 work also in the set-multilinear setting with nearly the same proofs. We sketch the ideas here, noting only the points of departure in the proofs. We assume that we are working with set-multilinear polynomials w.r.t. an equipartition $\mathcal{P} := \{X_1, \ldots, X_n\}$ into sets of size N.

For the second part of Theorem 3, we note that the proofs of Theorem 24, Theorem 25 and Theorem 26 above work almost without change in the set-multilinear case. For the first part of Theorem 3, we note that we can get something *stronger* in the set-multilinear setting by showing that if a set-multilinear ABP of size s computes a (p, ε)-almost full-rank polynomial, then there is a $(p, O(\log_N(sn/\varepsilon))$-balanced-chain set-system of size $O(s^2 n)$. The improvement is in the chain-balance parameter in the set-system where the base of the logarithm is now N instead of 2.

In the interesting case that $N = n^{\Theta(1)}$, the above paragraphs show that we get a tight connection between the smallest set-multilinear ABP computing a full-rank polynomial. For any c, there is a set-multilinear ABP of size $n^{O(c)}$ computing a full-rank polynomial if and only if there is an $O(c)$-balanced-chain set-system of size $n^{O(c)}$.

Finally, these results can be used to argue the lower bound for sums of interval-set-multilinear ABPs just as for the multilinear case.

References

1 Scott Aaronson. Multilinear formulas and skepticism of quantum computing. In László Babai, editor, *Proceedings of the 36th Annual ACM Symposium on Theory of Computing, Chicago, IL, USA, June 13-16, 2004*, pages 118–127. ACM, 2004. doi:10.1145/1007352.1007378.

2 Noga Alon, Mrinal Kumar, and Ben Lee Volk. Unbalancing sets and an almost quadratic lower bound for syntactically multilinear arithmetic circuits. *Comb.*, 40(2):149–178, 2020. doi:10.1007/S00493-019-4009-0.

3 C. S. Bhargav, Prateek Dwivedi, and Nitin Saxena. Lower bounds for the sum of small-size algebraic branching programs. *Theor. Comput. Sci.*, 1041:115214, 2025. doi:10.1016/J.TCS.2025.115214.

4 Prerona Chatterjee, Deepanshu Kush, Shubhangi Saraf, and Amir Shpilka. Lower bounds for set-multilinear branching programs. In Rahul Santhanam, editor, *39th Computational Complexity Conference, CCC 2024, July 22-25, 2024, Ann Arbor, MI, USA*, volume 300 of *LIPIcs*, pages 20:1–20:20. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2024. doi:10.4230/LIPIcs.CCC.2024.20.

5 Suryajith Chillara, Nutan Limaye, and Srikanth Srinivasan. Small-depth multilinear formula lower bounds for iterated matrix multiplication with applications. *SIAM J. Comput.*, 48(1):70–92, 2019. doi:10.1137/18M1191567.

6 E Csáki, P Erdős, and P Révész. On the length of the longest excursion. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 68(3):365–382, 1985.

7 Zeev Dvir, Guillaume Malod, Sylvain Perifel, and Amir Yehudayoff. Separating multilinear branching programs and formulas. In Howard J. Karloff and Toniann Pitassi, editors, *Proceedings of the 44th Symposium on Theory of Computing Conference, STOC 2012, New York, NY, USA, May 19 - 22, 2012*, pages 615–624. ACM, 2012. doi:10.1145/2213977.2214034.

8 Thomas Eiter and Heikki Mannila. Computing discrete fréchet distance. techreport cd-tr 94/64, 1994.

9 Michael A. Forbes, Amir Shpilka, Iddo Tzameret, and Avi Wigderson. Proof complexity lower bounds from algebraic circuit complexity. *Theory Comput.*, 17:1–88, 2021. doi:10.4086/TOC.2021.V017A010.

10 Deepanshu Kush and Shubhangi Saraf. Improved low-depth set-multilinear circuit lower bounds. In Shachar Lovett, editor, *37th Computational Complexity Conference, CCC 2022, July 20-23, 2022, Philadelphia, PA, USA*, volume 234 of *LIPIcs*, pages 38:1–38:16. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022. doi:10.4230/LIPIcs.CCC.2022.38.

11 Deepanshu Kush and Shubhangi Saraf. Near-optimal set-multilinear formula lower bounds. In Amnon Ta-Shma, editor, *38th Computational Complexity Conference, CCC 2023, July 17-20, 2023, Warwick, UK*, volume 264 of *LIPIcs*, pages 15:1–15:33. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023. doi:10.4230/LIPIcs.CCC.2023.15.

12 Noam Nisan. Lower bounds for non-commutative computation (extended abstract). In Cris Koutsougeras and Jeffrey Scott Vitter, editors, *Proceedings of the 23rd Annual ACM Symposium on Theory of Computing, May 5-8, 1991, New Orleans, Louisiana, USA*, pages 410–418. ACM, 1991. doi:10.1145/103418.103462.

13 Noam Nisan and Avi Wigderson. Lower bounds on arithmetic circuits via partial derivatives. *Comput. Complex.*, 6(3):217–234, 1997. doi:10.1007/BF01294256.

14 Ran Raz. Separation of multilinear circuit and formula size. *Theory Comput.*, 2(6):121–135, 2006. doi:10.4086/TOC.2006.V002A006.

15 Ran Raz. Multi-linear formulas for permanent and determinant are of super-polynomial size. *J. ACM*, 56(2):8:1–8:17, 2009. doi:10.1145/1502793.1502797.

16 Ran Raz, Amir Shpilka, and Amir Yehudayoff. A lower bound for the size of syntactically multilinear arithmetic circuits. *SIAM J. Comput.*, 38(4):1624–1647, 2008. doi:10.1137/070707932.

17 Ran Raz and Iddo Tzameret. The strength of multilinear proofs. *Comput. Complex.*, 17(3):407–457, 2008. doi:10.1007/S00037-008-0246-0.

18 Ran Raz and Amir Yehudayoff. Lower bounds and separations for constant depth multilinear circuits. *Comput. Complex.*, 18(2):171–207, 2009. doi:10.1007/S00037-009-0270-8.

19 Ran Raz and Amir Yehudayoff. Multilinear formulas, maximal-partition discrepancy and mixed-sources extractors. *J. Comput. Syst. Sci.*, 77(1):167–190, 2011. doi:10.1016/j.jcss.2010.06.013.