Electronic Colloquium on Computational Complexity, Revision 1 of Report No. 217 (2026)

Hitting sets From Derivations: Waring Rank and Noncommutative ABPs

STEWARDS*

Shubham Bhardwaj† Foram Lakhani‡

Partha Mukhopadhyay§ Ramprasad Saptharishi†

Pratik Shastri¶ Amit Kumar Sinhababu§

Abstract

We give an explicit hitting set of size $\text{poly}(n, d, r)$ for the class of n-variate degree-d polynomials whose partial derivative space is bounded by r, over any field $\mathbb{F}$ of characteristic zero. In particular, this yields a polynomial sized hitting set for the class of depth-3 powering circuits.

The proof extends to the setting of noncommutative ABPs and provides a polynomial sized hitting set (of matrices) for the class of noncommututive polynomials that can be computed by polynomial sized noncommutative ABPs (this was also one of the results reported by Open AI [Ope26] in their release on 6 Oct 2026).

The main technical insight is the construction of a "formal derivation" and properties of the associated Wronskian with respect to this derivation, which was previously studied by Moura [Mou04] in a very different context. The proofs in this paper are elementary and completely self-contained.

AI disclosure: A subset of the stewards obtained a weaker statement (hitting sets for polynomials with bounded partial derivative space) during conversations [Ast26] with OpenAI GPT-6 Astra. Extensions of this idea to the broader class of noncommutative ABPs were obtained by the other stewards (via further conversations with Astra), prior to OpenAI's dump on 6 Oct 2026.

*The origination of the ideas was via AI tools. We list ourselves as *stewards* to emphasise this, but take responsibility in providing an exposition of the external proof with the clarity and understanding that we believe would help any member of the community.

†Tata Institute of Fundamental Research, Mumbai, India. Email: {shubham.bhardwaj,ramprasad}@tifr.res.in.
Research supported by the Department of Atomic Energy, Government of India, under project number RTI400112, and Google and SERB Research Grants.

‡IIT Kanpur, India. Email: foramp23@iitk.ac.in.

§Chennai Mathematical Institute, India. Email: {partham,amitks}@cmi.ac.in

¶IT University of Copenhagen, Denmark. Email: prasha@itu.dk.

ISSN 1433-8092

The proofs presented in this writeup is a rewriting (in the authors' words) of the proofs for both settings obtained by the models in a form that we believe is understandable to researchers. Although the proof is technically identical to OpenAI's release, we believe this exposition is self-contained, simpler, and more valuable to the community. The absurdity of the situation is not lost on us.

1 Introduction

Algebraic circuits are the most natural computational model for studying the complexity of multivariate polynomials via the number of basic operations (additions, multiplications) required to compute it. The field of algebraic complexity broadly deals with the classification of polynomials based on their computational complexity, and also on algorithmic questions regarding polynomials provided via such computations.

Two of the most important aspects of this study are the question of lower bounds (i.e., finding explicit polynomials that cannot be computed by small algebraic circuits), and the algorithmic task of polynomial identity testing (i.e., given a circuit C, algorithmically check if C is computing the zero polynomial). In the context of polynomial identity tests, there are two types of algorithms — whitebox PITs (where the algorithm is given the entire circuit description), or blackbox PITs (where we are only provided evaluation access to the circuit). Although the questions of proving lower bounds and constructing deterministic PITs appear to be of very different flavours, they are intimately connected to each other. Several classical results [HS80, Agr05, KI04, CKS19, GKSS19] show that a strong enough lower bound for circuits can be used to obtain efficient deterministic polynomial identity tests, and vice-versa. Although similar statements when specialised to restricted circuit classes are known in some contexts [DSY09, Oli16, CKS19, BKR+26], this has broadly been a guiding principle for the study of lower bound and PITs for subclasses of circuits. Historically, a polynomial identity test for a class $\mathscr{C}$ has been a successor to a lower bound for the class $\mathscr{C}$, and is often heavily inspired by the techniques used in the lower bound proof. Paradoxically, even 'natural' "barriers" for proving algebraic circuit lower bounds stem from the conjectured existence of *succinct polynomial identity tests* [GKSS17, FSV18].

Over the recent years, there has been quite a lot of progress made towards proving lower bounds for natural subclasses of algebraic circuits, and also constructing whitebox and blackbox PITs for them [AGKS15, GKST15, GG20, LST21, AF22, GOS+26]. One subclass of circuits that has been just shy of a full resolution has been the class of *depth*-3 *powering circuits*.

Depth-3 powering circuits: A depth-3 powering circuit computes a polynomial of the form $f = c_1\ell_1^d + \cdots + c_s\ell_s^d$ where each ℓ_i is a linear polynomial over the underlying variables. This is the circuit analogue of the classical notion of Waring rank, where the Waring rank of a polynomial f is the smallest s such that f can be written as sum of s powers of linear forms. Saxena [Sax08]

gave the first lower bound for this class of circuits, and also a polynomial time white-box PIT. A different white-box PIT was also given by Kayal [Kay10]. Agrawal, Saha and Saxena [ASS13], and Forbes and Shpilka [FS13] gave the first $n^{O(\log n)}$-size hitting set for this class (among constructions of quasi-polynomial sized hitting sets for other more general classes), and Forbes, Saptharishi and Shpilka [FSS14] improved this to $n^{O(\log\log n)}$.

The key weakness of the above model that enabled all the lower bounds and PITs was the observation that any k-th order partial derivative of $f = c_1\ell_1^d + \cdots + c_s\ell_s^d$ is a linear combination of $\ell_1^{d-k}, \ldots, \ell_s^{d-k}$ and thus the space of all possible partial derivatives of f (which we shall denote by $\partial^{=*}(f)$) is only $O(sd)$-dimensional. The hitting sets constructed by [ASS13, FSS14] also extend to the (potentially) more general class of polynomials f whose space of all possible partial derivatives is low-dimensional. It continues to remain open if such polynomials with low-dimensional space of partial derivatives admit a not-too-large representation as a depth-3 powering circuit.

The first result of this paper is an explicit polynomial-sized hitting set for this class of polynomials over a characteristic zero field.

Theorem 1.1. *Let $\mathbb{F}$ be a field of characteristic zero. For any $n, r \geq 0$, there is an explicit polynomial map $(G_1(t), \ldots, G_n(t)) \in \mathbb{F}[t]^n$ with $\deg G_i \leq O(nr^2)$ such that if $f(x_1, \ldots, x_n)$ is a polynomial such that $\dim \partial^{=*} f \leq r$, then*

$$f(x_1, \ldots, x_n) = 0 \Leftrightarrow f(G_1(t), \ldots, G_n(t)) = 0.$$

In particular, the above yields a $\mathrm{poly}(n, d, s)$-sized hitting set for the class of size s, degree d depth-3 powering circuits. In fact, $G_i(t)$ can just be instantiated to be the *truncated Taylor series* of $\ln(1 + it)$:

$$G_i(t) = it - \frac{(it)^2}{2} + \frac{(it)^3}{3} - \cdots + (-1)^{B-1}\frac{(it)^B}{B} \quad \text{for } i = 1, \ldots, n$$

where $B = nr^2 + r$.

Using power series for constructing hitting-set generators: The general structure of constructing a hitting-set generator via power series is reminiscent of the randomised PIT algorithms of Chen and Kao [CK97], or Lewin and Vadhan [LV98]. We describe the setting of [CK97] (the setting of [LV98] is very similar with irreducible polynomial playing the role of primes). They show that if $f(x_1, \ldots, x_n)$ is a nonzero polynomial, then

$$f(\pm\sqrt{p_{1,1}} \pm \cdots \pm \sqrt{p_{1,\ell}}, \cdots, \pm\sqrt{p_{n,1}} \pm \cdots \pm \sqrt{p_{n,\ell}}) \neq 0$$

where $p_{i,j}$'s are distinct primes with $\ell = O(\log \deg(f))$. However, an efficient algorithm cannot evaluate f at these irrationals of infinite precision. [CK97] show that if the signs are randomly chosen, we can achieve an error of ε as long as we truncate the above irrationals to precision $B(\varepsilon)$. This results in randomised algorithm for PIT where reducing the error is achieved by increasing the precision (and thereby the running time) but keeping the number of random bits fixed.

The PITs in this paper is similar in the sense that the generator is appropriate truncations of infinite objects such as algebraically independent power series (e.g. $(P_i(t) := \log(1+it) \ : \ i \in [n])$) or linear operators in the noncommutative setting. For the class of polynomials with low-dimensional partial derivative space, we show that

$$0 \neq f \implies f(P_1(t), \ldots, P_n(t)) \neq 0 \bmod t^B$$

for a not-too-large B. For the noncommutative setting, the linear operators are closely related to the above, but are given by certain *iterated integrals*.

Mrinal Kumar pointed out to us that $f(e^{zt}, e^{z^2 t}, \ldots, e^{z^n t}) \neq 0 \bmod t^s$ for any nonzero s-sparse polynomial (thus giving a bivariate, low degree, hitting set generator for sparse polynomials). It would be interesting to see if this applies to a more general class of polynomials.

Noncommutative algebraic branching programs: Next, we will look at the PIT question for noncommutative models of computation. In particular, we look at noncommutative ABPs. Fix a field $\mathbb{F}$ and a set of variables $X = \{x_1, \ldots, x_n\}$. A noncommutative polynomial over $\mathbb{F}$ in X is a finite $\mathbb{F}$-linear combination of *words* in X^*. The set of all noncommutative polynomials in X forms a ring under addition and noncommutative products. It is denoted by $\mathbb{F}\langle X \rangle$.

Nisan, in his seminal paper [Nis91b], introduced the model of noncommutative Algebraic Branching Programs (also called noncommutative ABPs)[1]. This is a model for computing noncommutative polynomials. A noncommutative ABP is a layered directed acyclic graph with two distinguished vertices, a source and a sink. The first layer contains only the source and the last only the sink. Every edge goes from one layer to the next and is labeled by a homogeneous linear polynomial. The weight of a source to sink path is the (order-respecting) product of the edge labels on it, and the ABP is said to compute f. Nisan proved exponential lower bounds for this model [Nis91b]. The PIT problem for noncommutative ABPs has also been studied extensively. More than a decade after Nisan's work, Raz and Shpilka gave a deterministic polynomial-time whitebox PIT algorithm [RS05] for this problem. In the blackbox setting, a hitting set for noncommutative ABPs cannot have commuting entries, so one allows evaluations at matrices of polynomial dimension [FS13]. Forbes and Shpilka [FS13] studied the nc-ABP PIT problem in the black-box setting and gave a quasipolynomial size matrix hitting set.

[1] To the best of our knowledge, this was the first time algebraic branching programs were explicitly introduced as a model of computation, noncommutative or otherwise.

There has also been a long line of work extending the study of noncommutative computation to models allowing inversion gates, beginning with the work of Hrubeš and Wigderson [HW15]. In a breakthrough, a deterministic polynomial-time algorithm for noncommutative rational identity testing was obtained by Garg, Gurvits, Oliveira, and Wigderson [GGdOW16, GGOW20] using operator scaling. It was followed by algorithms based on purely algebraic techniques [IQS18], and using convex optimization [HH21]. In the black-box setting, randomized polynomial-time algorithm follows from the work of Derksen and Makam [DM17], while deterministic quasipolynomial-time hitting sets for noncommutative rational formulas were obtained in [ACM24]. Despite considerable effort, the Forbes-Shpilka algorithm has remained unbeaten for nc-ABP PIT, in the blackbox setting. The second result of this paper is an explicit polynomial size hitting set for the noncommutative ABPs.

Theorem 1.2. *Let $\mathbb{F}$ be a field of characteristic zero. There is a polynomial-time algorithm which, when given inputs $n, s \in \mathbb{N}$ in unary, outputs a tuple $T_1, \cdots, T_n$ of square matrices with rational entries such that for every nonzero noncommutative polynomial $f \in \mathbb{F}\langle x_1, \ldots, x_n\rangle$ computable by a homogeneous noncommutative ABP of size at most s, $f(T_1, \ldots, T_n) \neq 0$. The T_i's have dimension at most $ns^2 + s$.*

Extending the above result to the non-homogeneous case is straightforward using a standard argument. We provide brief details for completeness.

Corollary 1.3. *There is a polynomial-time algorithm which, given $n, s \in \mathbb{N}$ in unary, outputs an explicit hitting set $\mathscr{H}$ of at most s tuples of square matrices over $\mathbb{F}$, each of dimension $O(ns^6)$, such that for every nonzero polynomial $f \in \mathbb{F}\langle x_1, \ldots, x_n\rangle$ computable by a noncommutative ABP of size at most s, there exists $(A_1, \ldots, A_n) \in \mathscr{H}$ satisfying*

$$f(A_1, \ldots, A_n) \neq 0.$$

Proof. Let $f = f_0 + f_1 + \cdots + f_d$, where f_k denotes the homogeneous component of degree k. Since f is computed by an ABP of size at most s, we have $d \leq s$. By the standard homogenization construction, each f_k can be computed by a homogeneous ncABP of size at most $O(sd^2) = O(s^3)$. Let $S = O(s^3)$ fixed suitably. Applying Theorem 1.2 with parameters n, S, we obtain matrices

$$T_1, \ldots, T_n \in \mathbb{F}^{B \times B}, \qquad B = nS^2 + S = O(ns^6),$$

such that $f_k(T_1, \ldots, T_n) \neq 0$ whenever $f_k \neq 0$.

Now introduce a commuting indeterminate z and consider the matrix polynomial

$$\begin{aligned} F(z) &:= f(zT_1, \ldots, zT_n) \\ &= \sum_{k=0}^{d} z^k f_k(T_1, \ldots, T_n). \end{aligned}$$

By Theorem 1.2, its evaluation $f_k(T_1, \ldots, T_n)$ is nonzero. Thus $F(z)$ is a nonzero matrix polynomial of degree at most $d \leq s$. Consequently, there exists $z \in \{0, 1, \ldots, s\}$ such that $F(z) \neq 0$. Therefore, the set

$$\mathscr{H} := \{(zT_1, \ldots, zT_n) : z \in \{0, 1, \ldots, s\}\}$$

is a hitting set for all nonzero noncommutative ABPs of size at most s. □

1.1 Proof overview

It would be instructive to begin with the setting of polynomials with low-dimensional partial derivatives as it illustrates the core ideas, which also translates to the noncommutative ABP setting.

Let $f(x_1, \ldots, x_n)$ be some nonzero polynomial. We would like to understand if $f(P_1(t), \ldots, P_n(t))$ can be divisible by t^B for a large B (for if not, we can just truncate each $P_i(t)$ to B-terms, and we have a hitting set generator). Let us think of this substitution as an $\mathbb{F}(t)$-linear map (since this perspective would be useful for the noncommutative setting):

$$\Phi \colon \mathbb{F}(t)[x_1, \ldots, x_n] \to \mathbb{F}((t))$$
$$\Phi(g(t, x_1, \ldots, x_n)) = g(t, P_1(t), \ldots, P_n(t)).$$

By the chain rule of derivatives,

$$\frac{d}{dt}\Phi(g) = \frac{d}{dt}g(t, P_1, \ldots, P_n) = \Phi\left(\partial_t g + \sum_{i=1}^{n} \frac{dP_i}{dt} \cdot \partial_{x_i} g\right).$$

From the above, let us define a *derivation* $D : \mathbb{F}(t)[x_1, \ldots, x_n] \to \mathbb{F}((t))[x_1, \ldots, x_n]$ given by

$$D = \partial_t + \sum_{i=1}^{n} b_i(t) \cdot \partial_{x_i}$$

where $b_i(t) = \frac{dP_i}{dt}$ so that, for any $g \in \mathbb{F}(t)[x_1, \ldots, x_n]$, we have $\frac{d}{dt}\Phi(g) = \Phi(Dg)$. One particular derivation that would be extremely useful in this setting is the following:

$$D = \partial_t + \sum_{i=1}^{n} \frac{i}{1+it} \cdot \partial_{x_i}.$$

Derivations of similar shape was studied by [Mou04], and the main reason we would be using this is for the following miraculous property — if $h(t, x_1, \ldots, x_n) \in \mathbb{F}(t)[x_1, \ldots, x_n]$ with $Dh = 0$, then h is actually a scalar in $\mathbb{F}$. This property is sometimes referred to as D having *a trivial ring of constants*. For $b_i(t) = \frac{i}{1+it}$, the corresponding $P_i(t) = \int b_i(t)dt = \ln(1+it)$. The above property that $Dh = 0 \implies h \in \mathbb{F}$ can in fact be used to show that $\{\ln(1+it) : i \in [n]\}$ are algebraically

independent power series in t. Therefore, for any nonzero polynomial $f(x_1, \ldots, x_n)$, we have $f(P_1, \ldots, P_n)$ is a nonzero element of $\mathbb{F}[\![t]\!]$. We now want a way of arguing that if f was *simple*, then $f(P_1(t), \ldots, P_n(t)) \neq 0 \bmod t^B$ for a moderate B.

Now suppose the f has a low-dimensional partial derivative space, and say $\{f_1, \ldots, f_r\}$ is a basis of this space (with $f_r = f$). Since $f_1, \ldots, f_r$ are linearly independent polynomials, we also have that $\Phi(f_1), \ldots, \Phi(f_r)$ are linearly independent elements of $\mathbb{F}[\![t]\!]$. Therefore, the standard Wronskian $W(\Phi(f_1), \ldots, \Phi(f_r))$ is nonzero power series in t (although this is folklore, Bostan and Dumas [BD10] provide a clean proof of this statement). This Wronskian matrix (for which $W(\Phi(f_1), \ldots, \Phi(f_r))$ is the determinant of) is just Φ applied point-wise to the following D-Wronskian matrix $\mathscr{W}_D(f_1, \ldots, f_r)$:

$$\mathscr{W}_D(f_1, \ldots, f_r) = \begin{bmatrix} f_1 & f_2 & \cdots & f_r \\ D(f_1) & D(f_2) & \cdots & D(f_r) \\ \vdots & \vdots & \ddots & \vdots \\ D^{r-1}(f_1) & D^{r-1}(f_2) & \cdots & D^{r-1}(f_r) \end{bmatrix}.$$

Since the set $\{f_1, \ldots, f_r\}$ is closed under taking derivatives, one can show that the above matrix factorises as

$$\mathscr{W}_D(f_1, \ldots, f_r) = R(t) \cdot F$$

where $R(t)$ is an $r \times r$ matrix consisting entries that are rational functions in t (of moderate numerator and denominator degrees), and F is an upper-triangular matrix with 1's on the diagonal, with off-diagonal entries involving polynomials in x. A corollary of this structure is that $W_D(f_1, \ldots, f_r) = \det \mathscr{W}_D(f_1, \ldots, f_r)$ is just a rational function in t and does not depend on x at all! Hence, $W_D(f_1, \ldots, f_r) = W(\Phi(f_1), \ldots, \Phi(f_r)) = \frac{A(t)}{B(t)}$ where the degrees of A, B are not-too-large. If $\Phi(f) = 0 \bmod t^B$, every entry in the last column of $\mathscr{W}((\Phi(f_1), \ldots, \Phi(f_r)))$ is divisible by t^{B-r+1} and this would force $t^{B-r+1} \mid A(t)$. If B is suitably large, this would be impossible.

The noncommutative ABP setting: Turns out, the same proof works almost exactly along the same lines for the noncommutative ABP setting. Here, we need to define ∂_{x_i} as the "left-derivative" (i.e., for any word w we have $\partial_{x_i}(x_j w) = w$ if $j = i$, and zero otherwise). The same derivation operator defined above also satisfies the following — if $h(t, x_1, \ldots, x_n) \in \mathbb{F}(t)\langle x_1, \ldots, x_n \rangle$ is a noncommutative polynomial with $Dh = 0$, then $h \in \mathbb{F}$. That is, the derivation has a trivial ring of constants in the noncommutative setting as well.

We now wish to define an $\mathbb{F}(t)$-linear map $\Psi : \mathbb{F}(t)\langle x_1, \ldots, x_n \rangle \to \mathbb{F}(\!(t)\!)$ with the same property as earlier that $\frac{d}{dt}\Psi(f) = \Psi(Df)$. Analogous to earlier, we would be able to show that D having a

trivial ring of constants would translate to Ψ being injective. In other words, Ψ will map every noncommutative polynomial to a nonzero power series.

Let us normalise Ψ to ensure $\Psi(1) = 1$, and try to understand what power series Ψ should map a word w to. Certainly, for each $i \in [n]$, we need to ensure

$$\begin{aligned} &\frac{d}{dt}\Psi(x_i) = \Psi(Dx_i) = b_i(t)\Psi(1) = b_i(t) \\ \implies &\Psi(x_i) = \int b_i(t)dt = P_i(t) \end{aligned}$$

like earlier. Now, to understand what Ψ should map words of larger length to, consider $x_i w$ for some word w. If we want $\frac{d}{dt}\Psi(f) = \Psi(Df)$ for all f, then

$$\begin{aligned} &\frac{d}{dt}\Psi(x_i w) = \Psi(D(x_i w)) = b_i(t)\Psi(w) \\ \implies &\Psi(x_i w) = \int b_i(t)\Psi(w)dt \end{aligned}$$

which give the *iterated integrals* that we alluded to. In fact, the above can be thought of as replacing x_i by the linear operator T_i on the infinite dimensional vector space $\mathbb{F}[\![t]\!]$ that acts as

$$T_i(g(t)) = \text{“multiply } g(t) \text{ with } b_i(t) \text{ and integrate formally”}.$$

Now that we defined the injective map $\Psi : \mathbb{F}(t)\langle x_1, \ldots, x_n\rangle$ to $\mathbb{F}(\!(t)\!)$, let us focus on the polynomial for which we want a PIT. If f is computable by a size s noncommutative branching program C, then the partial derivative space (under left-derivatives) of f is in fact low-dimensional — a spanning set for this space is just $\{f_1, \ldots, f_s\}$ where f_i is the noncommutative polynomial computed by C between the i-th vertex and the sink. Once again, we wish to understand the Wronskian $W(\Psi(f_1), \cdots, \Psi(f_s))$, the associated matrix $\mathscr{W}(\Psi(f_1), \ldots, \Psi(f_s))$, and the related D-Wronskian $\mathscr{W}_D(f_1, \ldots, f_s)$. Along very similar lines, turns out we can factorise the matrix $\mathscr{W}_D(f_1, \ldots, f_s)$ as

$$\mathscr{W}_D(f_1, \ldots, f_s) = R(t) \cdot F$$

where $R(t)$ is once again a matrix of rational functions in t of moderate numerator and denominator degree, and F is an upper-triangular matrix with 1's on the diagonal and off-diagonal entries being elements of $\mathbb{F}\langle x_1, \ldots, x_n\rangle$. Hence, we get that $W(\Psi(f_1), \ldots, \Psi(f_s)) = \det R(t) = \frac{A(t)}{B(t)}$ of not-too-large degrees. Just like earlier, if $\Psi(f) = 0 \bmod t^B$, then we would be forced to admit that $t^{B-s+1} \mid A(t)$ and this is impossible if B is chosen large enough.

Therefore, we can truncate Ψ to just B terms of the power series, and this results in the linear operators T_i's being truncated to $B \times B$ matrices giving us the hitting set.

1.2 Classical sources of the key ideas

As is often the case with AI-generated proofs, the key ideas used in this paper have been discovered multiple times in completely different contexts. We try to give our account of them to the best of our understanding.

The notion of formal derivations opens into a general study of *derivation fields* and the specific derivation used is an example of something called Shamsuddin's derivations (see [Now94] for more details). The analysis of the order of vanishing via Wronskians can be traced to Moura [Mou03, Mou04]. The iterated integrals used in the noncommutative ABP setting are also understood as *hyperlogarithms*, or the Chen-Fliess series [Che54, Fli81] and their linear independence appears to have been studied extensively [Pan15, DDHS11]. The article by Espinosa et al [EEFGV26] provide some historical roots of Chen-Fliess series.

In this paper, we take the approach of giving a self-contained proof of the hitting set constructions in the style of a motivated exposition rather than merely citing the above papers, but we wish to stress that all of these ideas have classical roots in fields perhaps slightly removed from algebraic complexity. We give a version of the proof in this classical context in the Section A for readers who might be more familiar with the associated backgrounds.

Notation

- Throughout this paper we assume that we will be working with a field of characteristic zero. We use $\mathbb{F}[\![t]\!]$ to denote the ring of formal power series in t, and $\mathbb{F}(\!(t)\!)$ will denote the formal ring of Laurent series in t.
- We use $\mathbb{F}(t)\langle x_1, \ldots, x_n\rangle$ to denote the ring of polynomials on noncommuting variables $x_1, \ldots, x_n$ with coefficients from the field $\mathbb{F}(t)$. That is, elements of $\mathbb{F}(t)\langle x_1, \ldots, x_n\rangle$ are formal finite $\mathbb{F}(t)$-linear combination of *words* over the alphabet $\{x_1, \ldots, x_n\}$ (with ε denoting the empty word); $x_1x_2 - x_2x_1$ is thus a nonzero element of $\mathbb{F}(t)\langle x_1, \ldots, x_n\rangle$. Here, each x_i commutes with the coefficient field $\mathbb{F}(t)$.
- We use $\partial_{x_i}(f)$ as a shorthand for $\frac{\partial f}{\partial x_i}$. For $\mathbb{F}[x_1, \ldots, x_n]$, this would refer to standard partial derivatives. In the noncommutative setting of $\mathbb{F}\langle x_1, \ldots, x_n\rangle$, this will refer to *left-derivatives* (which shall be defined in due course).
- We abuse notation and sometimes use x to denote the tuple $(x_1, \ldots, x_n)$. Notations such as $\deg_x$ or $\mathrm{Hom}_x^{=d}$ will denote the degree or homogeneous component with respect to the variables $(x_1, \ldots, x_n)$.

2 Derivations and associated "evaluation" maps

The core of the PITs involve the following notion of *derivation operators*, and we define a special case that would be sufficient for this paper.

Definition 2.1 (Linear derivations)**.** *An operator* $D : \mathbb{F}(t)[x_1, \ldots, x_n] \to \mathbb{F}(t)[x_1, \ldots, x_n]$ *of the form*

$$Df = \partial_t f + \sum_{i=1}^{n} b_i(t) \cdot \partial_{x_i}(f)$$

where $b_i(t) \in \mathbb{F}(t)$ *shall be called a* linear derivation *on* $\mathbb{F}(t)[x_1, \ldots, x_n]$*. We shall call the set*

$$\ker D = \{h \in \mathbb{F}(t)[x_1, \ldots, x_n] \ : \ Dh = 0\}$$

as the ring of constants *of the derivation* D*. We shall say that* D *has a* trivial ring of constants *if* $\ker D = \mathbb{F}$. ◊

Associated with a derivation of the form $D = \partial_t + \sum_{i=1}^{n} b_i(t)\partial_{x_i}$, we can define an $\mathbb{F}(t)$-linear map $\Phi : \mathbb{F}(t)[x_1, \ldots, x_n] \to \mathbb{F}((t))$ that satisfies

$$\frac{d}{dt}\Phi(f) = \Phi(Df). \tag{2.2}$$

Indeed, if $\Phi(f) := f(t, P_1(t), \ldots, P_n(t))$ for some $P_i(t) \in \mathbb{F}[\![t]\!]$, then by the above we must have

$$\begin{aligned}\frac{d}{dt}\Phi(f) = \frac{d}{dt}f(t, P_1(t), \ldots, P_n(t)) &= \left(\partial_t f + \sum_i P_i'(t)\partial_{x_i} f\right)(t, P_1(t), \ldots, P_n(t)) \\ &= \Phi\left(\partial_t f + \sum_i P_i'(t)\partial_{x_i} f\right) = \Phi(Df)\end{aligned}$$

which forces $P_i'(t) = b_i(t)$. Thus, associated with the derivation D is the "evaluation map" that maps each x_i to $P_i(t)$ which is the formal integral of $b_i(t)$.

The following lemma shows that if D has a trivial ring of constants, then the map Φ is injective.

Lemma 2.3 (Injectivity of Φ)**.** *Suppose the linear derivation* D *has a trivial ring of constants. Then, the associated map* $\Phi\colon \mathbb{F}(t)[x_1, \ldots, x_n] \to \mathbb{F}((t))$ *defined above is injective.*

Proof. Suppose not. Consider the nonzero $f \in \mathbb{F}(t)[x_1, \ldots, x_n]$ with $\Phi(f) = 0$ that is minimal in the following sense:

- it has the smallest **x**-degree (say d),
- it has the fewest monomial of highest **x**-degree.

Without loss of generality (by scaling f by an element of $\mathbb{F}(t)$ if necessary), we may assume that one of the monomials of f of the highest x-degree has coefficient 1. Let $d = \deg_{\mathbf{x}}(f)$.

Since $\Phi(f) = 0$ we also have $\frac{d}{dt}\Phi(f) = \Phi(Df) = 0$. We claim that Df contradicts minimiality. This is because $Df = \partial_t f + \sum_i b_i(t)\partial_{x_i} f$, and $\sum_i b_i(t)\partial_{x_i} f$ has lower $\mathbf{x}$-degree, and $\partial_t f$ has the same x-degree but reduces the number of monomials of $\mathbf{x}$-degree d. Thus, by the minimality of the choice of f, we have $Df = 0$. However D has a trivial ring of constants which forces $f \in \mathbb{F}$ and this contradicts either $f \neq 0$ or $\Phi(f) = 0$. □

Therefore, if we want to find maps Φ that preserves nonzeroness, we are indirectly looking for suitable derivations D that have a trivial ring of constants.

2.1 A concrete derivation with a trivial ring of constants

There appears to be a rich literature on the study of *derivation fields*. with numerous examples of derivations that have a trivial ring of constants. The derivation we define below appears to be a special case of a broader class of derivations called Shamsuddin derivations. Nowicki's book [Now94] provides a fairly comprehensive study of this field.

For this paper, we will focus on the following derivation from here on.

$$D = \partial_t + \sum_{i=1}^{n} b_i(t)\partial_{x_i} \quad \text{where } b_i(t) = \frac{i}{1+it}.$$

Lemma 2.4 (D has a trivial ring of constants)**.** *The derivation D defined above has a trivial ring of constants. That is, if $h(t, x_1, \ldots, x_n) \in R = \mathbb{F}(t)[x_1, \ldots, x_n]$ such that $Dh = 0$, then $h \in \mathbb{F}$.*

Proof. Let h_i denote the degree i (in x) homogeneous component of h and let $\deg_x h = d$. If $d = 0$, we are done. Thus, we may assume that $d \geq 1$. Focusing on the degree d part of $D(h) = 0$, we have

$$\begin{aligned} 0 = \mathrm{Hom}_x^{=d}(D(h)) &= \partial_t h_d \\ \implies \partial_t h_d &= 0. \end{aligned}$$

Thus, the highest-degree homogeneous part of h is t-free. Now focusing on the homogeneous part of degree $d-1$, we have

$$\begin{aligned} 0 = \mathrm{Hom}_x^{=d-1}(D(h)) &= \left(\partial_t h_{d-1} + \sum_{i=1}^{n} \frac{i}{1+it} \cdot \partial_{x_i} h_d\right) \\ \implies \partial_t h_{d-1} &= -\sum_{i=1}^{n} \frac{i}{1+it} \cdot \partial_{x_i} h_d. \end{aligned}$$

Note that h_d is independent of t and thus so is $\partial_{x_i} h_d$. Hence, if $\partial_{x_i} h_d \neq 0$, then the RHS has a simple pole (of order 1) at $t = -\frac{1}{i}$ but $\partial_t h_{d-1}$ can never have simple poles. To see this, write the *partial fraction* decomposition of $h_{d-1} \in \mathbb{F}(t)[x_1, \ldots, x_n]$ as

$$h_{d-1}(t) = p(t, \mathbf{x})) + \sum_a \sum_{j \geq 1} \frac{c_{a,j}(\mathbf{x})}{(t-a)^j}$$

for some polynomial $p(t, \mathbf{x})$ and t-free $c_{a,j}(\mathbf{x}) \in \mathbb{F}[x_1, \ldots, x_n]$ for finitely many a. Differentiating term by term, $p(t, \mathbf{x})$ contributes a polynomial and thus no poles, while each term $c_{a,j}(\mathbf{x}) \cdot (t-a)^{-j}$ contributes $(-j) \cdot c_{a,j}(\mathbf{x}) \cdot (t-a)^{-j-1}$ and thus a pole of order $j+1$ at a. Hence, $\partial_t h_{d-1}$ has no pole of order exactly 1 at any point.) Thus, the only way the above can be zero is if $\partial_{x_i} h_d = 0$ for all $i \in [n]$, which is absurd when $d \geq 1$. Thus, $d = 0$ and hence $h \in \mathbb{F}$. □

Remark. *A sufficient condition for the above proof is to use $b_1(t), \ldots, b_n(t)$ such that each b_i has a distinct simple pole. The above example of $b_i(t) = \frac{i}{1+it}$ is one of the simplest example of this kind.* ♢

The evaluation map associated with the derivation D defined above is $\Phi : x_i \mapsto P_i(t)$ where $P_i(t) = \int b_i(t)dt = \ln(1+it)$. Thus, Lemma 2.3 and Lemma 2.4 show that $\{\ln(1+it) \ : \ i \in [n]\}$ is algebraically independent (in fact even over $\mathbb{F}(t)$).

3 Wronskians of derivative closed set of polynomials

Let $\{f_1, \ldots, f_r\}$ be a basis for $\partial^{=*}(f)$, and let us assume that $\deg f_1 \leq \cdots \leq \deg f_r$ with $f_1 = 1$ and $f_r = f$. Recall that $D = \partial_t + \sum_{i=1}^n b_i(t)\partial_{x_i}$ and the D-Wronskian of $f_1, \ldots, f_r$ is defined as

$$\mathscr{W}_D(f_1, \ldots, f_r) = \begin{bmatrix} f_1 & \cdots & f_r \\ Df_1 & \cdots & Df_r \\ \vdots & \ddots & \vdots \\ D^{r-1}f_1 & \cdots & D^{r-1}f_r \end{bmatrix}$$

$$W_D(f_1, \ldots, f_r) = \det \mathscr{W}_D(f_1, \ldots, f_r).$$

Lemma 3.1 (The D-Wronskian is $\mathbf{x}$-free)**.** *$W_D(f_1, \ldots, f_r)$ defined above, for any derivation of the form $D = \partial_t + \sum b_i(t)\partial_{x_i}$, is independent of $\mathbf{x}$.*

Proof. Since $\{f_1, \ldots, f_r\}$ is a basis for $\partial^{=*}(f)$, for each $i \in [n]$ and $j \in [r]$, note that

$$\begin{aligned} \partial_{x_i} f_j &= \sum_{k<j} a_{ijk} \cdot f_k \\ \implies \partial_{x_i} D^m f_j &= \sum_{k<j} a_{ijk} \cdot D^m f_k \quad \text{since } \partial_{x_i} Dg = D(\partial_{x_i} g) \text{ for any } g. \end{aligned}$$

Therefore, by the multilinearity of the determinant across columns, $\partial_{x_i} W_D(f_1, \ldots, f_r)$ is a sum of determinants of the matrix $\mathscr{W}_D(f_1, \ldots, f_r)$ with a column replaced by its partial derivative with respect to x_i. Thus, if C_j is the j-th column of $\mathscr{W}_D(f_1, \ldots, f_r)$, then

$$\partial_{x_i} C_j = \sum_{k<j} a_{ijk} \cdot C_k.$$

Therefore, $\partial_{x_i} W_D(f_1, \ldots, f_r) = 0$ for every $i \in [n]$. Hence, $W_D(f_1, \ldots, f_r)$ is independent of any of the $\mathbf{x}$-variables and is a function of t alone. □

The above lemma, along with properties of the derivation D defined in the previous section is sufficient to obtain a hitting set generator for polynomials with low partial derivative space. However, we can make a stronger observation that would be more useful when dealing with noncommutative ABPs in the later section.

Theorem 3.2. *There is an $r \times r$ matrix $F(x)$ and an $r \times r$ matrix $R(t)$ such that $\mathscr{W}_D = R(t)F(x)$. Moreover, the matrix $F(x)$ is upper triangular matrix with diagonal entries equal to* 1.

We can clearly see that Lemma 3.1 is an immediate consequence of this theorem.

Proof. Let $\mathbf{y} = \{y_1, y_2, \ldots, y_n\}$ be a set of new variables. Use Taylor series expansion to write (for $1 \leq j \leq r$)

$$f_j(\mathbf{y} + \mathbf{x}) = \sum_{\mathbf{e} \in \mathbb{N}^n} \frac{\mathbf{x}^{\mathbf{e}}}{\mathbf{e}!} \cdot \partial^{\mathbf{e}} f_j(y).$$

Using the derivative closed property we can further write

$$f_j(\mathbf{y} + \mathbf{x}) = f_j(\mathbf{y}) + \sum_{\ell<j} f_\ell(\mathbf{y}) \cdot F_{\ell,j}(\mathbf{x}), \tag{3.3}$$

for some polynomials $F_{\ell,j}(x)$. It is convenient to express these equations in matrix-vector notation. For that define $\widehat{f}(\mathbf{x}) = \begin{bmatrix} f_1(\mathbf{x}) & f_2(\mathbf{x}) & \cdots & f_r(\mathbf{x}) \end{bmatrix}$. Then from the equation above, it is straightforward to express

$$\widehat{f}(\mathbf{y} + \mathbf{x}) = \widehat{f}(\mathbf{y}) \cdot F(\mathbf{x}), \tag{3.4}$$

where $F(\mathbf{x})$ is the upper-triangular whose (ℓ, j)-th entry is just $F_{\ell,j}(\mathbf{x})$ from (3.3) with $F_{\ell,\ell} = 1$. Define $D_{\mathbf{y}} = \partial_t + \sum_{i=1}^n b_i(t)\partial_{y_i}$ (the analogue of D but with respect to $\mathbf{y}$), applying $D_{\mathbf{y}}$ to (3.4) (and using the fact that $F(\mathbf{x})$ is independent of t or $\mathbf{y}$), we get

$$D_{\mathbf{y}}\left(\widehat{f}(\mathbf{y} + \mathbf{x})\right) = D_{\mathbf{y}}\left(\widehat{f}(\mathbf{y}) \cdot F(\mathbf{x})\right) = D_{\mathbf{y}}\left(\widehat{f}(\mathbf{y})\right) \cdot F(\mathbf{x}).$$

Iterating the above,

$$D_{\mathbf{y}}^m\left(\widehat{f}(\mathbf{y}+\mathbf{x})\right) = D_{\mathbf{y}}^m\left(\widehat{f}(\mathbf{y})\right)\cdot F(\mathbf{x}) \quad \text{for all } m \geq 0.$$

In other words, if $W(t,\mathbf{x}) = W_D(f_1,\ldots,f_r)$, we have $W(t,\mathbf{x}+\mathbf{y}) = W(t,\mathbf{y})\cdot F(\mathbf{x})$. Setting $\mathbf{y}=0$, we get the factorisation $W(t,\mathbf{x}) = W(t,\mathbf{0})\cdot F(\mathbf{x})$. □

Corollary 3.5 (Degree bounds for $W_D(f_1,\ldots,f_r)$)**.** *For $D = \partial_t + \sum \frac{i}{1+it}\partial_{x_i}$, and $\{f_1,\ldots,f_r\}$ that is $\partial_{\mathbf{x}}$-closed, we have*

$$W_D(f_1,\ldots,f_r) = \frac{A(t)}{B(t)}$$

where $A, B \in \mathbb{F}[t]$ with $\deg A, \deg B \leq nr^2$.

Proof. Let $Q(t) = \prod_{i=1}^n (1+it)$. We prove by induction that for any $g(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$, we have $D^m g = \frac{A_m(t,x)}{Q^e}$ for some $A_m(t,x) \in \mathbb{F}[t,\mathbf{x}]$ with $\deg_t A_m \leq m(n-1)$ and $e = m$.

$$Dg = \partial_t g + \sum_{i=1}^n \frac{i}{1+it}\partial_{x_i} g = \frac{\sum_{i=1}^n (iQ(t)/(1+it))\cdot \partial_{x_i} g}{Q(t)} =: \frac{A_1(t,\mathbf{x})}{Q(t)^{e_1}}.$$

Clearly, $\deg_t A_1 \leq n-1$ and $e_1 = 1$. For induction,

$$\begin{aligned}
D^m g = D\left(\frac{A_{m-1}}{Q^{m-1}}\right) &= \partial_t\left(\frac{A_{m-1}}{Q^{m-1}}\right) + \frac{1}{Q^m}\cdot\left(\sum_{i=1}^n \frac{iQ}{(1+it)}\cdot\partial_{x_i}A_{m-1}\right)\\
&= \frac{\partial_t A_{m-1}}{Q^{m-1}} - \frac{eA_{m-1}\cdot\partial_t Q}{Q^m} + \frac{1}{Q^m}\cdot\left(\sum_{i=1}^n \frac{iQ}{(1+it)}\cdot\partial_{x_i}A_{m-1}\right)\\
&= \frac{Q\cdot\partial_t A_{m-1}}{Q^m} - \frac{eA_{m-1}\cdot\partial_t Q}{Q^m} + \frac{1}{Q^m}\cdot\left(\sum_{i=1}^n \frac{iQ}{(1+it)}\cdot\partial_{x_i}A_{m-1}\right)\\
&=: \frac{A_m(t,\mathbf{x})}{Q^m}
\end{aligned}$$

Thus, we get that $\deg_t A_m \leq \deg_t A_{m-1} + (n-1)$ and this completes the inductive step.

From the above bound on the t-degree of every entry of $\mathscr{W}_D(f_1,\ldots,f_r)$, and since Lemma 3.1 asserts its determinant is $\mathbf{x}$-free, we get that

$$W_D(f_1,\ldots,f_r) = \frac{A(t)}{B(t)}$$

for some $A, B \in \mathbb{F}[t]$ with $\deg A, \deg B \leq nr^2$. □

4 Hitting sets for polynomials with low partial derivative space

Now we have all the ingredients to prove Theorem 1.1. We state it here in a slightly different form from which Theorem 1.1 would readily follow.

Theorem 4.1. *Suppose* $0 \neq f(x_1, \ldots, x_n) \in \mathbb{F}[x_1, \ldots, x_n]$ *such that* $\dim \partial^{=*}(f) \leq r$. *Let* $P_1(t), \ldots, P_n(t) \in \mathbb{F}[\![t]\!]$ *be defined as*

$$P_i(t) = it - \frac{(it)^2}{2} + \frac{(it)^3}{3} - \frac{(it)^4}{4} + \cdots = \ln(1 + it)$$

Then, $f(P_1(t), \cdots, P_n(t)) \neq 0 \bmod t^B$ *for* $B = nr^2 + r$.

Theorem 1.1 follows from the above since we may truncate each P_i to the first B terms and retain the nonzeroness of f.

Proof of Theorem 4.1. For simplicity, let us assume that $\dim \partial^{=*} f = r$ (else we will replace r with the actual dimension in the proof below). Let $f_1, f_2, \ldots, f_r$ be a basis of the space V_f of all partial derivatives of f in increasing order of degree with $f_r = f$. Let $P(t) = (P_1(t), P_2(t), \ldots, P_n(t))$ where $P_i(t) = \ln(1 + it)$ for $i = 1, \ldots, n$ and let $\Phi : \mathbb{F}(t)[x_1, \ldots, x_n] \to \mathbb{F}(\!(t)\!)$ given by $\Phi(g(t, x_1, \ldots, x_n)) = g(t, P_1, \ldots, P_n)$.

By Lemma 2.3 and Lemma 2.4, we have that Φ is injective. Therefore, $\Phi(f_1), \ldots, \Phi(f_r)$ are linearly independent elements of $\mathbb{F}[\![t]\!]$. By [BD10], we have

$$W(\Phi(f_1), \ldots, \Phi(f_r)) = \det \mathscr{W}(\Phi(f_1), \ldots, \Phi(f_r)) \neq 0.$$

By since $\mathscr{W}(\Phi(f_1), \ldots, \Phi(f_r))$ is just the matrix $W_D(f_1, \ldots, f_r)$ with Φ applied to every entry, we get from Corollary 3.5 that

$$W(\Phi(f_1), \ldots, \Phi(f_r)) = \frac{A(t)}{B(t)}$$

for some nonzero polynomials A, B of degree at most nr^2.

Now suppose $\Phi(f) = 0 \bmod t^B$ for $B = nr^2 + r$, we have $\frac{d^m}{dt^m}\Phi(f) = 0 \bmod t^{B-m}$ for all $0 \leq m \leq r - 1$. Thus, the entire last column of $\mathscr{W}(\Phi(f_1), \ldots, \Phi(f_r))$ is divisible by t^{B-r+1} and hence we must have that t^{B-r+1} divides $A(t)$. But this is impossible since $\deg A \leq nr^2 < B - r + 1$.

Thus, we must have that $\Phi(f) = f(P_1(t), \ldots, P_n(t)) \neq 0 \bmod t^{nr^2+r}$.

□

5 Extensions to noncommutative ABPs

The approach for polynomials with low-partial derivative space extends to noncommutative ABPs as well. The ring we are operating on is $\mathbb{F}(t)\langle x_1, \ldots, x_n\rangle$. We will be using the same derivation[2] D on this ring, defined by

$$D = \partial_t + \sum_{i=1}^{n} b_i(t) \cdot \partial_{x_i}$$

but here ∂_{x_i} refers to the *left-derivative* with respect to x_i. That is, for any word w, define

$$\partial_{x_i}(x_j w) = \begin{cases} w & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

We can prove the analogue of Lemma 2.4 even in the context of $R = \mathbb{F}(t)\langle x_1, \ldots, x_n\rangle$.

Lemma 5.1 (D has a trivial ring of constants, noncommutatively)**.** *The derivation $D : \mathbb{F}(t)\langle x_1, \ldots, x_n\rangle \to \mathbb{F}(t)\langle x_1, \ldots, x_n\rangle$ defined above has a trivial ring of constants. That is, if $h \in R$ such that $Dh = 0$, then $h \in \mathbb{F}$.*

Proof. As earlier, let h_i denote the degree i (in x) homogeneous component of h and say $\deg_{\mathbf{x}} h = d$. As in the proof of Lemma 2.4, since $D(h) = 0$ we can conclude that h_d is t-free.

For any word w, let us use c_w to denote the coefficient of the word w in h (which is an element of $\mathbb{F}(t)$). If we take any word w of length $d-1$, then

$$\begin{aligned} 0 = \operatorname{coeff}_w(Dh) &= \partial_t c_w + \sum_{i=1}^{n} b_i(t) \cdot c_{x_i w} \\ \implies -\partial_t c_w &= \sum_{i=1}^{n} \frac{i}{1+it} \cdot c_{x_i w}. \end{aligned}$$

Once again, the LHS cannot have any simple poles and hence $c_{x_i w} = 0$ for all $i \in [n]$. Since w is arbitrary, this forces $h_d = 0$ which is absurd. Thus, $d = 0$ and hence $h \in \mathbb{F}$. □

5.1 The "evaluation" map related to D in the noncommutative setting

As earlier, associated with the above D, we wish to define an $\mathbb{F}(t)$-linear map $\Psi : \mathbb{F}(t)\langle x_1, \ldots, x_n\rangle \to \mathbb{F}(\!(t)\!)$ such that for any $f \in R = \mathbb{F}(t)\langle x_1, \ldots, x_n\rangle$

$$\frac{d}{dt}\Psi(f) = \Psi(Df).$$

[2]Strictly speaking, the above map D is not a '*derivation*' in the traditional sense since it does not satisfy the Leibniz rule (i.e., $D(fg) \neq fD(g) + D(f)g$) due to the non-commutativity. Nevertheless, we abuse this terminology and refer to it as a 'derivation' for the sake of exposition.

Unlike the commutative case, this is not merely a substitution $x_i \mapsto P_i(t)$ extended to all words since then $\Psi(x_i x_j - x_j x_i) = 0$ and thus Ψ is not injective on R. Therefore, we need to define Ψ on each word w. Again, if $\Psi(1) = 1$, then we are forced to define Ψ inductively on larger words as

$$\begin{aligned} \frac{d}{dt}\Psi(x_i w) &= \Psi(D(x_i w)) \\ &= \Psi(b_i(t)w) = b_i(t)\Psi(w) \\ \implies \Psi(x_i w) &= \int b_i(t)\Psi(w)dt. \end{aligned}$$

Thus, if T_i is the linear operator on $\mathbb{F}[\![t]\!] \to \mathbb{F}[\![t]\!]$ defined as

$$T_i \colon g(t) \mapsto \int b_i(t)g(t)dt$$

then the above map Ψ can be thought of as substituting in place of each x_i the operator T_i.

Exactly as in Lemma 2.3, we now can show that Ψ defined above is injective.

Lemma 5.2 (Ψ is injective)**.** *Suppose the linear derivation D has a trivial ring of constants in $R = \mathbb{F}(t)\langle x_1, \ldots, x_n\rangle$. Then, the associated map $\Psi \colon R \to \mathbb{F}(\!(t)\!)$ defined above is injective.*

Proof. Follows the proof of Lemma 2.3 verbatim. □

Remark. *An important distinction here from the commutative case is that the map Φ defined there was infact an algebra homomorphism. However, the map here is just an $\mathbb{F}(t)$-linear map. This is of course necessary for injectivity as any algebra homomorphism from $\mathbb{F}(t)\langle x_1, \ldots, x_n\rangle$ to $\mathbb{F}(\!(t)\!)$ must necessarily satisfy $\Psi(x_1x_2) = \Psi(x_2x_1)$ and hence cannot be injective. Nevertheless, the map Ψ constructed about will eventually be thought of as an algebra homomorphism to the space of* linear maps on $\mathbb{F}[\![t]\!]$ *(instead of to $\mathbb{F}[\![t]\!]$).* ◊

5.2 Noncommutative ABPs have low-dimensional partial derivative space

Definition 5.3 (Noncommutative Algebraic Branching Programs (ncABPs) [Nis91a])**.** *A* noncommutative algebraic branching program (ncABP) *of size s is given by a directed acyclic graph of s vertices $\{v_1, \ldots, v_s\}$ with directed edges (v_i, v_j) with $i < j$ having weight ℓ_{ij} that is a linear polynomial in $x_1, \ldots, x_n$.*

The polynomial computed by the ncABP is the sum, over all directed paths from v_1 to v_s, of the product of the edge weights. Concretely,

$$f(x_1, \ldots, x_n) = \sum_{v_{i_0}=v_1, v_{i_1}, \cdots, v_{ir}=v_s} \prod_{j=0}^{t-1} \ell_{i_j, i_{j+1}}$$

More generally, we can define $f_{ij}(x_1, \ldots, x_n)$ analogous as the polynomial computed between vertex v_i and v_j by a similar sum over all directed paths from v_i to v_j.

Such an ncABP is said to be homogeneous *of degree d if the underlying directed acyclic graph is layered into $d+1$ layers with v_1 being the unique vertex in layer 1 and v_s being the unique vertex in layer $d+1$, and all edges only going from vertices in layer i to layer $i+1$ with all edge-weights ℓ_{ij} being linear forms in $x_1, \ldots, x_n$ (i.e., ℓ_{ij} has no constant term).* ◊

It is well-known [Nis91a, RS04] that if f is a degree d polynomial that is computed by an ncABP of size s, then each homogeneous component of f can be computed by an ncABP of size at most $O(sd^2)$. Furthermore for the purposes of PIT, it suffices to focus on just homogeneous polynomials computable by a homogeneous ABP. See [RS04, FS13] for details. Thus, from here on, we will assume that f is a homogeneous degree d polynomial that is computable by a homogeneous ncABP of size s.

The following lemma is implicit in the work of Raz and Shpilka [RS04].

Lemma 5.4 (ncABPs are derivative closed [RS04])**.** *Suppose f is computable by a noncommutative branching program of size s. If $\partial_{\mathbf{x}}^{=*}(f)$ refers to span of all left-derivatives of f of all order, then $\partial^{=*}(f)$ is contained in the span of $\{f_{i,s} : i \in [s]\}$. Furthermore, if the ncABP is minimal (in the sense that there is no smaller ABP for f), then $\{f_{i,s} : i \in [s]\}$ is a linearly independent set. Similarly, if the ncABP is minimal and homogeneous, then $f_{1,1}, \ldots, f_{1,s}$ are linearly independent.*

5.3 Factorisation of the Wronskian matrix

Let $f \in \mathbb{F}\langle X \rangle$ have degree d, and consider an ABP for f. For each $i \in [n]$, let M_i be the *adjacency matrix* of the ABP with respect to variable x_i. Concretely, $(M_i)_{j,k}$ is the coefficient of the variable x_i in the edge weight $\ell_{j,k}$. Note that M_i is a strictly upper-triangular matrix. Define F as the $s \times s$ matrix where $F_{ij} = f_{i,j}(x_1, \ldots, x_n)$, the polynomial computed between vertex v_i and v_j. In particular F is an upper-triangular matrix with all diagonal entries being 1 by definition. Let $M = \sum_i M_i x_i$ be the adjacency matrix of the ABP.

Observation 5.5. $F = I + M + M^2 + \cdots + M^d$

Proof. This follows from the definition of F, combined with the fact that M is the adjacency matrix of the (layered) ABP. □

Observation 5.6. *If $\partial_{x_i} F$ refers to the matrix obtained by applying the left derivative ∂_{x_i} to every entry of F, then*

$$\partial_{x_i} F = M_i F.$$

Proof. From Observation 5.5, we have

$$F = I + (M_1x_1 + \cdots + M_nx_n) + \cdots + (M_1x_1 + \cdots + M_nx_n)^d$$

Thus, $\partial_{x_i} F = M_i + M_i(M_1x_1 + \cdots + M_nx_n) + \cdots + M_i(M_1x_1 + \cdots + M_nx_n)^{d-1}$. Note that $M_i(M_1x_1 + \cdots + M_nx_n)^d = 0$: every coefficient in the expansion is 0 as these are coefficients of the monomials of degree $d+1$ with leftmost variable x_i, but no such monomial exists in the ABP. Therefore, $\partial_{x_i} F = M_i F$ as claimed. □

Let us use f_i to denote $f_{1,i}$ defined above. Without loss of generality (by taking the smallest ncABP for f), we may assume that $\{f_1, \ldots, f_s\}$ are linearly independent.

Recall the derivation $D = \partial_t + \sum_{i=1}^n b_i(t)\partial_{x_i}$ and the D-Wronskian matrix

$$\mathscr{W}_D(f_1, \ldots, f_s) = \begin{bmatrix} f_1 & f_2 & \cdots & f_s \\ Df_1 & Df_2 & \cdots & Df_s \\ \vdots & \vdots & \ddots & \vdots \\ D^{s-1}f_1 & D^{s-1}f_2 & \cdots & D^{s-1}f_s \end{bmatrix}.$$

Lemma 5.7 (Factorisation of $\mathscr{W}_D(f_1, \ldots, f_s)$)**.** *There is an $s \times s$ matrix R consisting of rational functions entries in t such that $\mathscr{W}_D(f_1, \ldots, f_s) = R \cdot F$.*

Proof. We will construct R one row at a time and show that for all $0 \leq m < s$

$$\begin{bmatrix} D^m f_1 & \cdots & D^m f_s \end{bmatrix} = R_m F$$

for suitably chosen row vector R_m. We would then have our factorisation $\mathscr{W}_D(f_1, \ldots, f_s) = R \cdot F$ as required.

For $m = 0$, we have $R_0 = [1\ 0\ \cdots\ 0]$. Suppose for some $m \geq 0$ we have

$$\begin{aligned} \begin{bmatrix} D^m f_1 & \cdots & D^m f_s \end{bmatrix} &= R_m F \\ \implies \begin{bmatrix} D^{m+1} f_1 & \cdots & D^{m+1} f_s \end{bmatrix} &= D\left(\begin{bmatrix} D^m f_1 & \cdots & D^m f_s \end{bmatrix}\right) \\ &= D(R_m \cdot F) = \partial_t(R_m \cdot F) + \sum_i b_i(t) \cdot \partial_{x_i}(R_m \cdot F) \\ &= (\partial_t R_m + R_m B)F \quad \text{where } B = \sum_{i=1}^n b_i(t) M_i \\ &=: R_{m+1} \cdot F \end{aligned}$$

Thus, $R_{m+1} = \partial_t R_m + R_m B$, and stacking the R_i's into a matrix R gives us the claim. □

Unlike the commutative case, defining the notion of a determinant of the above matrix needs

care as the entries involve noncommuting variables. But the above factorisation is enough to show that the ordinary Wronskian $W(\Psi(f_1), \ldots, \Psi(f_s))$ is a rational function with low-degree numerator and denominator.

Corollary 5.8. *Let $\mathscr{W}(t) := \mathscr{W}(\Psi(f_1), \ldots, \Psi(f_s))$ be the ordinary Wronskian of $\Psi(f_1), \ldots, \Psi(f_s)$. Then $\mathscr{W}(t) = R(t)\Psi(F)$, where $\Psi(F)$ is the matrix obtained from F by applying the map Ψ entry-wise. As a result* $\det(\mathscr{W}(t)) = \det(R(t))$.

Proof. By Lemma 5.7, $\mathscr{W}_D(f_1, \ldots, f_s)_{a,b} = \sum_{c=1}^{s} R(t)_{a,c} F_{c,b}$. Since Ψ is $\mathbb{F}(t)$-linear, applying it entry-wise gives

$$\Psi(\mathscr{W}_D(f_1, \ldots, f_s)_{a,b}) = \sum_{c=1}^{s} R(t)_{a,c} \Psi(F_{c,b}).$$

This also implies that

$$\Psi(\mathscr{W}_D(f_1, \ldots, f_s)) = R(t)\Psi(F). \tag{5.9}$$

However for $1 \leq i \leq s$, $\frac{d}{dt}\Psi(f_i) = \Psi(Df_i)$. Differentiating m times, we get that

$$\frac{d^m}{dt^m}\Psi(f_i) = \Psi(D^m f_i).$$

Combining this with Eq. (5.9), we get that $\mathscr{W}(t) = R(t)\Psi(F)$. Since F is an upper triangular matrix with diagonal entries 1, so is $\Psi(F)$, and it follows that $\det(\mathscr{W}(t)) = \det(R(t))$. □

Corollary 5.10 (Degree bounds for the entries of R)**.** *If $D = \partial_t + \sum_{i=1}^{n} \frac{i}{1+it} \cdot \partial_{x_i}$, then $\det R(t)$ is a rational function of the form $\frac{P(t)}{Q(t)}$ with $\deg P, Q \leq ns^2$.*

Proof. Just as in Corollary 3.5, using the recurrence for the rows of R that $R_{i+1} = R_i' + R_i B$, each entry can be shown to be a rational function with numerator and denominator of degree at most ns. Since R is an $s \times s$ matrix, we get the required degree bound. □

5.4 Obtaining the hitting set generator

Lemma 5.11 (Bounding the order of $\Psi(f)$)**.** *Let $f(x_1, \ldots, x_n) \in \mathbb{F}\langle x_1, \ldots, x_n \rangle$ be a nonzero polynomial that is computable by an ncABP of size at most s. Let $\Psi \colon \mathbb{F}(t)\langle x_1, \ldots, x_n \rangle \to \mathbb{F}(\!(t)\!)$ as defined above. Then, $\Psi(f) \neq 0 \bmod t^B$ for $B = ns^2 + s$.*

Proof. Assume on the contrary that $\Psi(f) = 0 \bmod t^B$ for $B = ns^2 + s$. Then, for each $m < s$ we have $\frac{d^m}{dt^m}\Psi(f) = 0 \bmod t^{ns^2+1}$. Let span $\{f_{1,1}, \ldots, f_{1,s}\}$ have as a basis $\{f_1, \ldots, f_s\}$ (assuming without loss of generality that s is the minimal ncABP size for f) with $f_s = f$.

By Corollary 5.8 and Corollary 5.10, we have that

$$W(\Psi(f_1),\ldots,\Psi(f_s)) = \det \mathscr{W}(\Psi(f_1),\ldots,\Psi(f_s)) = \frac{P(t)}{Q(t)}$$

with $\deg P, Q \leq ns^2$. Furthermore, since Ψ is an injective map (Lemma 5.2), we know that the above Wronskian determinant is nonzero [BD10]. On the other hand, the last column of $\mathscr{W}(\Psi(f_1),\ldots,\Psi(f_s))$ is divisible by t^{ns^2+1} which is impossible since $\deg P(t) \leq ns^2$. □

The final step is to convert the above statement into suitable matrices that create the hitting set.

Theorem 5.12. *There is an explicit set of matrices $\tilde{T}_1,\ldots,\tilde{T}_n \in \mathbb{F}^{B\times B}$, for $B = ns^2 + s$ such that any nonzero noncommutative polynomial $f(x_1,\ldots,x_n) \in \mathbb{F}\langle x_1,\ldots,x_n\rangle$ that is computable by a homogeneous ncABP of size at most s satisfies $f(\tilde{T}_1,\ldots,\tilde{T}_n) \neq 0$.*

Proof. Consider the infinite linear operators $T_1,\ldots,T_n\colon \mathbb{F}[\![t]\!] \to [\![t]\!]$ given by

$$\begin{aligned} T_i : g(t) = g_0 + g_1 t + \cdots \mapsto & \int b_i(t) g(t) dt \\ &= \int i(1 - it + (it)^2 + \cdots) g(t) dt \\ &= h_0 t + h_1 \frac{t^2}{2} + h_2 \frac{t^3}{3} + \cdots \\ \text{where } h(t) = h_0 + h_1 t + \cdots &= i(1 - it + (it)^2 - (it)^3) \cdot g(t). \end{aligned}$$

Each $\tilde{T}_i$ is just a truncation of T_i to $\mathbb{F}[t]/t^B$ and thus is an $B \times B$ matrix.

Recall that $\Psi : \mathbb{F}(t)\langle x_1,\ldots,x_n\rangle \to \mathbb{F}(\!(t)\!)$ maps words over $\{x_1,\ldots,x_n\}$ to elements of $\mathbb{F}[\![t]\!]$. For length 1 words, x_i, we have $\Psi(x_i) = \int b_i(t)dt = T_i(1)$. More generally for any word $x_{i_1}\cdots x_{i_r}$ we inductively defined

$$\begin{aligned} \Psi(x_{i_1}\cdots x_{i_r}) &= \int b_{i_1}(t)\Psi(x_{i_2}\cdots x_{i_r})dt \\ &= T_{i_1}\left(T_{i_2}\cdots T_{i_r}(1)\right) = T_{i_1}\cdots T_{i_r}(1). \end{aligned}$$

Thus, Ψ maps each word $x_{i_1}\cdots x_{i_r}$ to the power series in t that is the image of the composition $T_{i_1}\cdots T_{i_r}$. Hence, $\Psi(f) = f(T_1,\ldots,T_n)(1)$. Since from Lemma 5.11 we have $\Psi(f) \neq 0 \bmod t^B$, we get that $f(\tilde{T}_1,\cdots,\tilde{T}_n)(1) \neq 0$ and hence $f(\tilde{T}_1,\cdots,\tilde{T}_n)$ is a nonzero linear map and thus a nonzero $B \times B$ matrix. □

Acknowledgements: We'd like to thank several members in the community (especially Robert Andrews, Umang Bhaskar, Prerona Chatterjee, Prahladh Harsha, Mrinal Kumar, Amir Shpilka and Anamay Tengse) for numerous philophical discussions on whether or not to publish this.

References

[ACM24] V. Arvind, Abhranil Chatterjee, and Partha Mukhopadhyay. Black-Box Identity Testing of Noncommutative Rational Formulas in Deterministic Quasipolynomial Time. In *STOC*, pages 106–117, 2024.

[AF22] Robert Andrews and Michael A. Forbes. Ideals, determinants, and straightening: proving and using lower bounds for polynomial ideals. In *STOC '22: 54th Annual ACM SIGACT Symposium on Theory of Computing, Rome, Italy, June 20 - 24, 2022*, pages 389–402. ACM, 2022.

[AGKS15] Manindra Agrawal, Rohit Gurjar, Arpita Korwar, and Nitin Saxena. Hitting-Sets for ROABP and Sum of Set-Multilinear Circuits. *SIAM Journal of Computing*, 44(3):669–697, 2015. Pre-print available at `arXiv:1406.7535`.

[Agr05] Manindra Agrawal. Proving Lower Bounds Via Pseudo-random Generators. In *Proceedings of the 25th International Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2005)*, volume 3821 of *Lecture Notes in Computer Science*, pages 92–105. Springer, 2005.

[ASS13] Manindra Agrawal, Chandan Saha, and Nitin Saxena. Quasi-polynomial hitting-set for set-depth-Δ formulas. In *Proceedings of the 45th Annual ACM Symposium on Theory of Computing (STOC 2013)*, pages 321–330, 2013. `eccc:TR12-113`.

[Ast26] OpenAI GPT 6 Astra. Building a polynomial hitting set, 2026.

[BD10] Alin Bostan and Philippe Dumas. Wronskians and Linear Independence. *Am. Math. Mon.*, 117(8):722–727, 2010.

[BKR+26] Somnath Bhattacharjee, Mrinal Kumar, Shanthanu S. Rai, Varun Ramanathan, Ramprasad Saptharishi, and Shubhangi Saraf. Closure under Factorization from a Result of Furstenberg. In *Proceedings of the 58th Annual ACM Symposium on Theory of Computing, STOC 2026, Salt Lake City, UT, USA, June 22-26, 2026*, pages 174–185. ACM, 2026.

[BR11] Jean Berstel and Christophe Reutenauer. *Noncommutative Rational Series with Applications*, volume 137 of *Encyclopedia of Mathematics and Its Applications*. Cambridge University Press, 2011.

[Che54] Kuo-Tsai Chen. Iterated Integrals and Exponential Homomorphisms. *Proceedings of the London Mathematical Society*, 4:502–512, 1954.

[CK97] Zhi-Zhong Chen and Ming-Yang Kao. Reducing Randomness via Irrational Numbers. In *Proceedings of the 29th Annual ACM Symposium on Theory of Computing (STOC 1997)*, pages 200–209, 1997.

[CKS19] Chi-Ning Chou, Mrinal Kumar, and Noam Solomon. Closure Results for Polynomial Factorization. *Theory of Computing*, 15(13):1–34, 2019.

[DDHS11] Matthieu Deneufchâtel, Gérard H. E. Duchamp, Vincel Hoang Ngoc Minh, and Allan I. Solomon. Independence of Hyperlogarithms over Function Fields via Algebraic Combinatorics. In *Algebraic Informatics: 4th International Conference, CAI 2011, Linz, Austria, June 21–24, 2011, Proceedings*, volume 6742 of *Lecture Notes in Computer Science*, pages 127–139, Berlin, Heidelberg, 2011. Springer. Pre-print available at `arXiv:1101.4497`.

[DM17] Harm Derksen and Visu Makam. Polynomial Degree Bounds for Matrix Semi-invariants. *Advances in Mathematics*, 310:44–63, 2017.

[DSY09] Zeev Dvir, Amir Shpilka, and Amir Yehudayoff. Hardness-Randomness Tradeoffs for Bounded Depth Arithmetic Circuits. *SIAM J. Comput.*, 39(4):1279–1293, 2009.

[EEFGV26] Luis A Duffaut Espinosa, Kurusch Ebrahimi-Fard, W Steven Gray, and GS Venkatesh. What is a Chen–Fliess Series? *Notices of the American Mathematical Society*, 73(2), 2026.

[Fli81] Michel Fliess. Fonctionnelles causales non linéaires et indéterminées non commutatives. *Bulletin de la Société Mathématique de France*, 109:3–40, 1981.

[FS13] Michael A. Forbes and Amir Shpilka. Quasipolynomial-Time Identity Testing of Non-commutative and Read-Once Oblivious Algebraic Branching Programs. In *Proceedings of the 54th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2013)*, pages 243–252, 2013. Full version at `arXiv:1209.2408`.

[FSS14] Michael A. Forbes, Ramprasad Saptharishi, and Amir Shpilka. Hitting sets for multilinear read-once algebraic branching programs, in any order. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing (STOC 2014)*, pages 867–875, 2014.

[FSV18] Michael A. Forbes, Amir Shpilka, and Ben Lee Volk. Succinct Hitting Sets and Barriers to Proving Lower Bounds for Algebraic Circuits. *Theory of Computing*, 14(1):1–45, 2018.

[GG20] Zeyu Guo and Rohit Gurjar. Improved Explicit Hitting-Sets for ROABPs. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2020, Virtual Conference, August 17-19, 2020*, volume 176 of *LIPIcs*, pages 4:1–4:16. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020.

[GGdOW16] Ankit Garg, Leonid Gurvits, Rafael Mendes de Oliveira, and Avi Wigderson. A Deterministic Polynomial Time Algorithm for Non-commutative Rational Identity Testing. In *FOCS*, pages 109–117, 2016.

[GGOW20] Ankit Garg, Leonid Gurvits, Rafael Oliveira, and Avi Wigderson. Operator Scaling: Theory and Applications. *Foundations of Computational Mathematics*, 20(2):223–290, 2020.

[GKSS17] Joshua A. Grochow, Mrinal Kumar, Michael E. Saks, and Shubhangi Saraf. Towards an algebraic natural proofs barrier via polynomial identity testing. *CoRR*, abs/1701.01717, 2017. Pre-print available at `arXiv:1701.01717`.

[GKSS19] Zeyu Guo, Mrinal Kumar, Ramprasad Saptharishi, and Noam Solomon. Derandomization from Algebraic Hardness: Treading the Borders. In *60th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2019, Baltimore, Maryland, USA, November 9-12, 2019*, pages 147–157. IEEE Computer Society, 2019.

[GKST15] Rohit Gurjar, Arpita Korwar, Nitin Saxena, and Thomas Thierauf. Deterministic Identity Testing for Sum of Read-once Oblivious Arithmetic Branching Programs. In *Proceedings of the 30th Annual Computational Complexity Conference (CCC 2015)*, pages 323–346, 2015. `arXiv:1411.7341`.

[GOS+26] Abhibhav Garg, Rafael Oliveira, Akash Kumar Sengupta, Nir Shalmon, and Amir Shpilka. Rank Bounds and Polynomial-Time PIT for Σ^k Π Σ Π^2 Circuits. In *41st Computational Complexity Conference, CCC 2026, Lisbon, Portugal, August 3-6, 2026*, volume 383 of *LIPIcs*, pages 17:1–17:18. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2026.

[HH21] Masaki Hamada and Hiroshi Hirai. Computing the nc-Rank via Discrete Convex Optimization on CAT(0) Spaces. *SIAM Journal on Applied Algebra and Geometry*, 5(3):455–478, 2021.

[HS80] Joos Heintz and Claus-Peter Schnorr. Testing Polynomials which Are Easy to Compute (Extended Abstract). In *Proceedings of the 12th Annual ACM Symposium on Theory of Computing (STOC 1980)*, pages 262–272. ACM, 1980.

[HW15] Pavel Hrubeš and Avi Wigderson. Non-commutative Arithmetic Circuits with Division. *Theory of Computing*, 11(14):357–393, 2015.

[IQS18] Gábor Ivanyos, Youming Qiao, and K. V. Subrahmanyam. Constructive Non-commutative Rank Computation Is in Deterministic Polynomial Time. *Computational Complexity*, 27(4):561–593, 2018.

[Kay10] Neeraj Kayal. Algorithms for Arithmetic Circuits. *Electronic Colloquium on Computational Complexity (ECCC)*, 2010.

[KI04] Valentine Kabanets and Russell Impagliazzo. Derandomizing Polynomial Identity Tests Means Proving Circuit Lower Bounds. *Computational Complexity*, 13(1-2):1–46, 2004. Preliminary version in the *35th Annual ACM Symposium on Theory of Computing (STOC 2003)*.

[LST21] Nutan Limaye, Srikanth Srinivasan, and Sébastien Tavenas. Superpolynomial Lower Bounds Against Low-Depth Algebraic Circuits. In *Proceedings of the 62nd Annual IEEE Symposium on Foundations of Computer Science (FOCS 2021)*, pages 804–814. IEEE, 2021. Preliminary version in the Electronic Colloquium on Computational Complexity (ECCC), Technical Report TR21-081.

[LV98] Daniel Lewin and Salil P. Vadhan. Checking Polynomial Identities over any Field: Towards a Derandomization? In *Proceedings of the 30th Annual ACM Symposium on Theory of Computing (STOC 1998)*, pages 438–447, 1998.

[Mou03] Claire Moura. Bounds on the vanishing order for the solutions of a linear differential system. *Journal of Dynamical and Control Systems*, 9(1):73–88, 2003.

[Mou04] Claire Moura. On the multiplicity of hyperelliptic integrals. *Nonlinearity*, 17(6):2057–2068, August 2004.

[Nis91a] Noam Nisan. Lower bounds for non-commutative computation. In *Proceedings of the 23rd Annual ACM Symposium on Theory of Computing (STOC 1991)*, pages 410–418, 1991. Available on `citeseer:10.1.1.17.5067`.

[Nis91b] Noam Nisan. Lower bounds for non-commutative computation (Extended Abstract). In *Proceedings of the Twenty-Third Annual ACM Symposium on Theory of Computing (STOC '91)*, pages 410–418, New Orleans, Louisiana, USA, 1991. ACM.

[Now94] Andrzej Nowicki. *Polynomial derivations and their rings of constants*. 10 1994.

[Oli16] Rafael Oliveira. Factors of low individual degree polynomials. *Comput. Complex.*, 25(2):507–561, 2016.

[Ope26] OpenAI. One Rational Matrix Hitting Point for Noncommutative Formulas. OpenAI Math Release preprint OAI:One-Rational-Matrix-Hitting-Point-for-Noncommutative-Formulas-September-24-2026, 2026.

[Pan15] Erik Panzer. *Feynman Integrals and Hyperlogarithms*. PhD thesis, Humboldt-Universität zu Berlin, 2015. See Chapter 3, Lemma 3.3.5.

[RS04] Ran Raz and Amir Shpilka. Deterministic Polynomial Identity Testing in Non-Commutative Models. In *Proceedings of the 19th Annual IEEE Conference on Computational Complexity (CCC 2004)*, pages 215–222, 2004.

[RS05] Ran Raz and Amir Shpilka. Deterministic polynomial identity testing in non-commutative models. *Computational Complexity*, 14(1):1–19, 2005. Preliminary version in the *19th Annual IEEE Conference on Computational Complexity (CCC 2004)*.

[Sax08] Nitin Saxena. Diagonal Circuit Identity Testing and Lower Bounds. In *Proceedings of the 35th International Colloquium on Automata, Languages and Programming (ICALP 2008)*, pages 60–71, 2008. Pre-print available at `eccc:TR07-124`.

[Sch61] Marcel-Paul Schützenberger. On the Definition of a Family of Automata. *Information and Control*, 4(2–3):245–270, 1961.

A Tracing back the classical methods

In this section, we provide a version of OpenAI's proof [Ope26] of black-box hitting point for noncommutative ABPs in the language of the literature it utilizes.

Let $\mathbb{F}\langle x_1, \ldots, x_n\rangle$ be the free associative algebra. Given a polynomial $f \in \mathbb{F}\langle x_1, \ldots, x_n\rangle$ computed by an ABP of size s, it is a standard fact that f has an s-dimensional linear representation consisting of $u, v \in \mathbb{F}^s$, and jointly nilpotent matrices $A_i \in \mathbb{F}^{s\times s}$ such that

$$f = \sum_w (u^T A_w v) w,$$

where $w = x_{i_1} \ldots x_{i_k}$ and $A_w = A_{i_1} \ldots A_{i_k}$, $A_\emptyset = I_s$. Moreover, $A_w = 0$ for $|w| > s$. This follows from the standard linear-representation point of view of recognizable series [BR11, Sch61].

For $1 \leq i \leq n$, define

$$b_i(t) = \frac{i}{1-it}, \qquad (T_i h)(t) = \int_0^t b_i(z)h(z)dz,$$

where the integration is coefficient-wise for $h(t) \in \mathbb{F}[[t]]$. For every word define the iterated integrals as follows

$$I_\emptyset = 1, \qquad I_{x_i w}(t) = (T_i I_w)(t).$$

The linear map $\Psi : \mathbb{F}\langle x_1, \ldots, x_n \rangle \to \mathbb{F}[[t]]$ is defined as

$$\Psi(w) = I_w(t),$$

and extending linearly. The map Ψ is well-studied in the context of nonlinear control theory and $\Psi(f)$ is known as Chen-Fliess series [Che54, Fli81]. The following result is also well-known.

Lemma A.1. *[Pan15, DDHS11]* $f \neq 0 \Rightarrow \Psi(f) \neq 0$.

Now construct the matrix-values formal series

$$Z(t) = \sum_w I_w(t) A_w \in (\mathbb{F}[[t]])^{s \times s}.$$

Differentiating with respect to t and using the easy fact that

$$I'_{x_j v}(t) = b_j(t) I_v(t),$$

we get the following

$$Z'(t) = \sum_{j,v} b_j I_v A_j A_v.$$

Arranging the terms further

$$Z'(t) = \left(\sum_{j=1}^n b_j A_j \right) \left(\sum_v I_v(t) A_v \right) = A(t) Z(t), \tag{A.2}$$

where $A(t) = \sum_{i=1}^n \frac{iA_i}{1-it}$. Let $y(t) = Z(t)v$ and

$$g(t) = u^T y(t) = \sum_w (u^T A_w v) I_w(t) = \Psi(f)(t).$$

By Lemma A.1, $g(t) \neq 0$. Moreover using Eq. (A.2),

$$y'(t) = Z'(t)v = A(t)y(t). \tag{A.3}$$

Note that the rational matrix $A(t)$ has no pole at $t = 0$. Given t_0 and a Taylor series expansion

$$h(t) = \sum_{k=0}^{\infty} c_k (t - t_0)^k,$$

define

$$\operatorname{ord}_{t_0} h = \min\{k \geq 0 : c_k \neq 0\}.$$

Our goal is to prove that $\operatorname{ord}_0 g \leq N$ for a polynomially bounded integer N. For that we use a classical result by Moura [Mou04, Theorem 1] which is also attributed to [Mou03]. We state the result in a form suitable for our purpose. Moura states the result over $\mathbb{C}$, but it easily extends to arbitrary fields of characteristic zero.

Theorem A.4. *[Mou04, Theorem 1] Suppose $y'(t) = A(t)y(t)$ be an s-dimensional rational linear system, t_0 is not a pole of $A(t)$, and $g(t) = u^T y(t)$ is not identically zero. Let P be the sum of the order of all poles (including infinity) of the differential form $\Omega = A(t)dt$. Then,*

$$\operatorname{ord}_{t_0} g \leq s - 1 + \binom{s}{2}(P - 2).$$

We note a few standard facts.

1. Ordinary point: t_0 is an ordinary point when each rational entry of $A(t)$ is defined at t_0.
2. Pole at a finite point a: the pole order of $A(t)dt$ at a is the largest pole order among the rational functions in $A(t)$.
3. Pole at infinity: substitute $z = 1/t$. Then $dt = -z^{-2}dz$. The order of the pole at infinity is the order of the pole at $z = 0$.

We apply the above to our case. For us,

$$A(t) = \sum_{i=1}^{n} \frac{iA_i}{1 - it}.$$

Let $\Omega = A(t)dt$. There are n finite poles of order 1 for $A(t)$. Substitute $t = \frac{1}{z}$. Then,

$$A(t)dt = -\left(\sum_{i=1}^{n} \frac{iA_i}{z - i}\right) z^{-1}dz.$$

So the order of the pole at infinity is at most 1. Hence $P \leq n+1$. Applying Theorem A.4,

$$\operatorname{ord}_0 g \leq s - 1 + \binom{s}{2}(n-1).$$

Set,

$$N \coloneqq s - 1 + \binom{s}{2}(n-1).$$

Truncation: Each $T_i : \mathbb{F}[[t]] \to \mathbb{F}[[t]]$ is a linear operator. Define the vector space $V_N = \mathbb{F}[[t]]/(t^{N+1})$. Naturally, each T_i induces an $\mathbb{F}$-linear map

$$T_i' : V_N \to V_N, \qquad T_i'(h) = T_i(h) \mod t^{N+1}.$$

By direct expansion

$$T_i'(t^k) = \sum_{r=k+1}^{N} \frac{i^{r-k}}{r} t^r.$$

Hence, relative to the basis $(1, t, \ldots, t^N)$ the matrix T_i' can be represented as

$$(T_i')_{r,k} = \begin{cases} \dfrac{i^{r-k}}{r}, & 0 \leq k < r \leq N, \\ 0, & \text{otherwise.} \end{cases} \tag{A.5}$$

Now we are ready to state the main theorem.

Theorem A.6. *For every nonzero noncommutative polynomial $f \in \mathbb{F}\langle x_1, \ldots, x_n \rangle$ computable by an ABP of size at most s, $f(T_1', \ldots, T_n') \neq 0$.*

Proof. Using the connection with iterated integrals, the main thing to notice is that

$$f(T_1', \ldots, T_n')1 = \sum_w c_w I_w(t) = \Psi(f)(t) \mod t^{N+1}.$$

However, we already proved that $\operatorname{ord}_0 g(t) \leq N$ where

$$g(t) = \sum_w (u^T A_w v) I_w(t).$$

Hence $f(T_1', \ldots, T_n')1 = g(t) \mod t^{N+1} \neq 0$ and the proof of the theorem follows. □

ECCC ISSN 1433-8092
https://eccc.weizmann.ac.il