Electronic Colloquium on Computational Complexity, Report No. 166 (2026)

Quasi-polynomial Frege Simulation of IPS beyond Noncommutativity

Abhranil Chatterjee* Prerona Chatterjee† Utsab Ghosal‡ Partha Mukhopadhyay§

September 4, 2026

Abstract

Grochow and Pitassi (2018) introduced the algebraic proof system, the Ideal Proof System (IPS), which connects algebraic circuit complexity to propositional proof complexity. They showed that propositional proof systems such as Extended Frege (Frege) are equivalent to circuit IPS (formula IPS) if the correctness of PIT for circuits (formulas) can be proved in Extended Frege (Frege). Along the same lines, the work of Li, Tzameret, and Wang (2018) gives a quasi-polynomial (unconditional) equivalence between noncommutative formula IPS and Frege (over $\mathbf{GF}(2)$) by proving the correctness of the noncommutative formula (ABP) PIT algorithm of Raz and Shpilka (2005) in Frege.

We extend this line of work beyond the fully noncommutative setting. We consider a partially commutative model in which the variables are partitioned into two sets: variables within each set are noncommuting while they commute across the sets. We show that the same quasi-polynomial equivalence with Frege continues to hold in this setting, even when both sets have unbounded size.

Enroute to the main result, we give a new deterministic polynomial time PIT algorithm for partially commutative formulas (ABPs) over such mixed variables and show that the correctness of the algorithm can be efficiently verified in Frege. To the best of our knowledge, this is the strongest known class of algebraic formulas for which Frege efficiently proves identity testing. Our approach uses tools from skew field theory (Cohn 1995), which may be of independent interest in algebraic proof complexity.

*Department of Computer Science and Engineering, IIT Kharagpur. Email: abhranilc@cse.iitkgp.ac.in

†School of Computer Sciences, National Institute of Science Education and Research, Bhubaneswar, India; Homi Bhabha National Institute, Anushakti Nagar, Mumbai, India. Email: prerona.ch@gmail.com

‡Imperial College, London. This work was done as a PhD student at Chennai Mathematical Institute, India. Email: utsab.ghosal@gmail.com.

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

ISSN 1433-8092

1 Introduction

The theory of propositional proof complexity studies the computational resources required to prove propositional tautologies in concrete proof systems. Cook and Reckhow [CR79] formalized the notion of a propositional proof system and showed that there exists a polynomially bounded proof system if and only if NP = coNP. Among the standard propositional proof systems, the *Frege proof system* plays a central role. A Frege proof[1] starts from a fixed finite collection of axiom schemes and derives new Boolean formulas using a fixed collection of sound inference rules. By the work of Reckhow [Rec76], the complexity of proving propositional tautologies in this system is independent of the particular choice of axioms and rules. Thus, Frege gives a robust model for measuring the complexity of ordinary propositional reasoning.

Despite being one of the main strong propositional proof systems, the power of Frege is still poorly understood. In particular, no super-polynomial lower bound is known for unrestricted Frege proofs; see, for example [Jeř25]. However, strong lower bounds are known for restricted versions such as bounded-depth Frege [Ajt88, PBI93, KPW95, PRST16, HR25]. Moreover, several families of tautologies that are hard for weaker proof systems admit polynomial-size Frege proofs. One such classical example is the pigeonhole principle [Bus87]. Thus, understanding the power and limitations of unrestricted Frege remains a basic problem in proof complexity.

This gap makes it useful to look for other ways to understand the power of Frege. One such direction is through algebraic proof systems. In an algebraic proof system, propositional formulas are translated into polynomial equations and proofs are represented by algebraic computations. This brings tools from algebraic complexity into proof complexity and gives a direct way to compare the complexity of proofs with the complexity of algebraic circuits; see, for example, [PT16] for a survey.

A central system in this direction is the Ideal Proof System (IPS), introduced by Grochow and Pitassi [GP18]. An IPS refutation is a single algebraic circuit (formula, ABP, etc.), that witnesses a given system of polynomial equations is unsatisfiable. The circuit version of IPS polynomially simulates Extended Frege, while the formula version polynomially simulates Frege [GP18]. At the same time, lower bounds for unrestricted IPS are closely connected to lower bounds in algebraic circuit complexity. Therefore, understanding the relation between Frege and algebraic proof systems such as IPS gives another approach to understanding the power of Frege.

Commutative IPS. We first recall the commutative Ideal Proof System introduced by Grochow and Pitassi [GP18].

> Let $\mathbb{F}$ be a field and let $f_1, \ldots, f_m \in \mathbb{F}[X]$ be a set of unsatisfiable polynomials, where the Boolean axioms $x_i^2 - x_i$ are included among the f_is. A commutative IPS certificate

[1] The formal definition is given in Subsection 2.3.

is a polynomial $P(X, Z) \in \mathbb{F}[X, Z]$ such that

$$P(X, \bar{0}) = 0 \qquad \text{and} \qquad P(X, f_1(X), \ldots, f_m(X)) = 1.$$

The size of a refutation is the size of an algebraic circuit[2] computing $P(X, Z)$. We call the proof system $\mathcal{C}$-IPS, if the proof polynomial $P(X, Z)$ is computed in $\mathcal{C}$. The proof system is sound by definition, and completeness follows from Hilbert's weak Nullstellensatz. Note that if the proof polynomial $P(X, Z)$ is computed by an algebraic class $\mathcal{C}$, then the verification of $\mathcal{C}$-IPS is exactly polynomial identity testing in $\mathcal{C}$.

A natural question is whether strong propositional proof systems can efficiently prove the soundness of IPS[3]. This question is fundamental because, as is standard in proof complexity, if a proof system P_2 admits polynomial-size proofs of the soundness of another proof system P_1, then P_2 polynomially simulates P_1. Therefore polynomial-size proofs of the soundness of IPS in Frege (or Extended Frege) immediately yield polynomial simulations of IPS by these propositional proof systems.

Grochow and Pitassi [GP18] approached this question in two steps. They first introduced a propositional formula $\mathsf{Soundness}_{\mathsf{IPS}}$ expressing the soundness of commutative IPS, and showed that polynomial-size proofs of a natural collection of PIT principles imply polynomial-size proofs of $\mathsf{Soundness}_{\mathsf{IPS}}$. They then proved that polynomial-size proofs of $\mathsf{Soundness}_{\mathsf{IPS}}$ imply polynomial simulations of commutative IPS by the ambient propositional proof system. Consequently, proving the soundness of commutative IPS reduces to proving the correctness of Polynomial Identity Testing (PIT) within propositional proof systems. Subsequent work further clarified this connection. Grochow [Gro26] showed that commutative IPS can be p-simulated by a Cook–Reckhow proof system [CR79] if and only if PIT $\in$ NP. Hence, understanding how propositional proof systems reason about polynomial identities becomes a central question in relating algebraic and propositional proof systems. The simulation theorem of Grochow and Pitassi reduces the simulation of commutative IPS by propositional proof systems to proving the correctness of Polynomial Identity Testing within those proof systems. At present, no strong propositional proof system is known to admit polynomial-size proofs of these PIT principles in the commutative setting.

Our question is more specific. Rather than asking whether unrestricted IPS can be simulated by some Cook–Reckhow proof system, we ask for which natural algebraic restrictions of IPS the simulation can be carried out by *Frege itself*. Equivalently, we want to understand the classes of algebraic identities that Frege can efficiently prove to establish the soundness of the corresponding IPS system. In this sense, we seek to understand the boundary of algebraic reasoning that Frege can efficiently simulate.

[2] or sometimes an algebraic formula. See Definition 2.1 for the definitions of algebraic circuits and formulas.

[3] Given $\{f_1, \ldots, f_m\}$, if there is a proof polynomial $P(X, Z)$ satisfying these two properties from the definition, then indeed the set has no common root.

Noncommutative IPS**.** Li, Tzameret, and Wang [LTW18] identified an important point on this boundary by considering the fully noncommutative setting. Instead of reasoning about the commutative PIT principles, they directly proved the soundness of a noncommutative variant of IPS (over $\mathbf{GF}(2)$) within Frege. This yields a quasi-polynomial simulation of noncommutative IPS by Frege.

They introduced the noncommutative Ideal Proof System (NC-IPS) as follows.

> Let $\mathbb{F}$ be a field and let $f_1, \ldots, f_m \in \mathbb{F}\langle x_1, \ldots, x_n\rangle$[4] be noncommutative unsatisfiable polynomials containing the Boolean axioms $\{x_i(1-x_i) : i \in [n]\}$ and the commutator axioms $\{x_i x_j - x_j x_i : i < j\}$.
>
> A noncommutative IPS certificate is a noncommutative polynomial $P(X, Z) \in \mathbb{F}\langle X, Z\rangle$ satisfying
>
> $$P(X, \bar{0}) = 0 \qquad \text{and} \qquad P(X, f_1, \ldots, f_m) = 1.$$
>
> The size of an NC-IPS refutation is the size of a noncommutative formula computing $P(X, Z)$.

The proof of Li, Tzameret, and Wang relies on a *reflection principle* for noncommutative IPS. Informally, such a principle states that Frege can prove the soundness of noncommutative IPS. As discussed above, proving the soundness of one proof system within another immediately yields a simulation. Thus it suffices to establish polynomial-size Frege proofs of the soundness of noncommutative IPS.

To prove this reflection principle, Li, Tzameret, and Wang reduce the problem to proving Boolean tautologies arising from noncommutative polynomial identities. Over the field $\mathbf{GF}(2)$, every noncommutative arithmetic formula F naturally defines a Boolean formula $\widetilde{F}$ by replacing addition with XOR and multiplication with AND. Thus every noncommutative polynomial identity F gives rise to a Boolean tautology $\neg\widetilde{F}$.

Using the deterministic Polynomial Identity Testing algorithm of Raz and Shpilka [RS05], they proved that every homogeneous noncommutative polynomial identity of size s has a corresponding Boolean tautology with a Frege proof of size $\mathrm{poly}(s)$. This theorem allows Frege to verify the polynomial identities appearing in a noncommutative IPS certificate. Consequently, Frege proves the soundness of noncommutative IPS, and therefore quasi-polynomially[5] simulates noncommutative IPS. The result of Li, Tzameret, and Wang illustrates a broader phenomenon. The existence of an efficient algorithm for Polynomial Identity Testing does not, by itself, imply an efficient Frege proof of the corresponding polynomial identities. In general, it is unclear whether the correctness of an efficient algorithm can be proved within Frege with only polynomial overhead. Thus, obtaining Frege simulations requires more than efficient algorithms. One also needs

[4] $\mathbb{F}\langle x_1, \ldots, x_n\rangle$ is the ring of noncommutative polynomials. Here the variables do not commute. That is, $\forall\ i \neq j\ x_i x_j - x_j x_i \neq 0$.

[5] The source of quasi-polynomial blow up is homogenization of the noncommutative formulas.

algorithms whose correctness admits efficient Frege proofs. The deterministic PIT algorithm of Raz and Shpilka [RS05] has precisely this property for noncommutative formulas.

The result of Li, Tzameret, and Wang therefore leaves a natural gap. At one endpoint, fully noncommutative formula IPS is already quasi-polynomially equivalent to Frege. At the other endpoint, in the fully commutative setting, an efficient simulation of formula IPS by Frege remains tied to proving suitable commutative PIT statements inside Frege. This suggests studying intermediate algebraic models in which some commutativity is introduced while asking whether the Frege simulation continues to hold.

Our motivation here is not that introducing commutativity should make Frege lower bounds easier. In fact, allowing more commutativity makes the underlying algebraic model stronger and can make algebraic lower bounds more difficult. Rather, we use these intermediate models to study how far the ability of Frege to efficiently reason about algebraic identities extends beyond the fully noncommutative setting. In particular, we ask:

> *How much commutativity can be introduced into the noncommutative* IPS *model while retaining an efficient simulation by Frege? More generally, what is the largest class of algebraic identities whose corresponding Boolean tautologies admit efficient Frege proofs?*

Answering this question would identify new algebraic proof systems whose soundness can be proved within Frege and hence admit efficient Frege simulations.

1.1 Our Contribution

Partially Commutative IPS. Motivated by the gap between commutative and noncommutative IPS, we define a family of proof systems called *Partially Commutative* IPS, $\{\mathsf{PC}_{m,n}\text{-}\mathsf{IPS}\}_{1\le m\le n}$. It bridges noncommutative IPS and Commutative IPS in the following sense: $\mathsf{PC}_{1,n}\text{-}\mathsf{IPS} \equiv \mathsf{NC}\text{-}\mathsf{IPS}$ and $\mathsf{PC}_{n,n}\text{-}\mathsf{IPS} \equiv \mathsf{IPS}$ where n is the total number of variables.

We begin by looking at the notion of a *partially commutative polynomial*. A partially commutative polynomial is defined over variables partitioned into buckets. Variables inside the same bucket do not commute, whereas variables belonging to different buckets commute. We explain it with a 2-partition of the variable set below.

> Let X_1, X_2 be two sets of variables such that $X_1 = \{x_{1,1}, \ldots, x_{1,n_1}\}$ and $X_2 = \{x_{2,1}, \ldots, x_{2,n_2}\}$, and let $\mathbb{F}$ be a field. A partially commutative monomial $m \in \mathbb{F}\langle X_1 \sqcup X_2\rangle$ is uniquely determined by its restrictions to X_1 and to X_2: $m|_{X_1} = m_{X_1} \in \mathbb{F}\langle X_1\rangle$ and $m|_{X_2} = m_{X_2} \in \mathbb{F}\langle X_2\rangle$, where m_{X_1} and m_{X_2} are noncommutative monomials. All variables in X_1 commute with all variables in X_2. Hence, for $m, m' \in \mathbb{F}\langle X_1 \sqcup X_2\rangle$, we have $m = m'$ if and only if $m|_{X_1} = m'|_{X_1}$ and $m|_{X_2} = m'|_{X_2}$. This induces an equivalence relation $\sim$ on the set of monomials.[6]

[6]For example, consider two monomials $m_1 = x_{1,1}x_{2,1}x_{1,2}x_{2,2}$ and $m_2 = x_{2,1}x_{1,1}x_{2,2}x_{1,2}$. These two monomials are the same with $m_1|_{X_1} = m_2|_{X_1} = x_{1,1}x_{1,2}$ and $m_1|_{X_2} = m_2|_{X_2} = x_{2,1}x_{2,2}$.

A partially commutative polynomial $f \in \mathbb{F}\langle X_1 \sqcup X_2 \rangle$ is an $\mathbb{F}$-linear combination of the partially commutative monomials over X_1 and X_2. A partially commutative polynomial over p buckets, say $f \in \mathbb{F}\langle X_1 \sqcup \cdots \sqcup X_p \rangle$, can be defined analogously.

This is a very well-studied model with connections to the problem of multiplicity equivalence testing of multi-tape automata [HK91, Wor13]. The study of one-way multi-tape finite automata was initiated in the seminal paper by Rabin and Scott [RS59]. The multiplicity equivalence testing problem for multi-tape automata is reduced to the identity testing problem over the partially commutative ring, and a deterministic polynomial-time algorithm was designed only very recently (over $\mathbb{Q}$) when the number of tapes is $O(1)$ [ACM24].

In particular, the bucketing structure of the variables defines a *noncommuting graph*, a structure (motivated by the classical work of Cartier and Foata [CF69]) which measures the amount of noncommutativity present in the model.

The noncommuting graph has one vertex for each variable in the polynomial, and an edge exists between two vertices if and only if the corresponding variables do not commute.

Note that a commutative polynomial $f \in \mathbb{F}[X]$ can be thought of as a partially commutative polynomial in $\mathbb{F}\langle X_1 \sqcup \cdots \sqcup X_n \rangle$ where $|X_i| = 1$. Thus, a natural generalization of both IPS and NC-IPS is the proof system $\mathsf{PC}_{m,n}$-IPS, which we define in such a way that $\mathsf{PC}_{1,n}$-IPS $\equiv$ NC-IPS and $\mathsf{PC}_{n,n}$-IPS $\equiv$ IPS where n is the total number of variables.

In particular, the input equations of a $\mathsf{PC}_{m,n}$-IPS proof are partially commutative polynomials in $\mathbb{F}\langle X_1 \sqcup \cdots \sqcup X_m \rangle$. Using this partially commutative ring, we define the corresponding Ideal Proof System. The definition is analogous to NC-IPS, except that computations are carried out over the partially commutative ring. A formal definition of $\mathsf{PC}_{m,n}$-IPS is given below.

Definition 1.1. (Partially commutative IPS ($\mathsf{PC}_{m,n}$-IPS)) *Let $\mathbb{F}$ be a field and $X = X_1 \sqcup \cdots \sqcup X_m$ be a set of partially commutative variables*[7] *with $|X| = n$ and for every $i \in [m]$, $|X_i| = n_i$.*

Assume that $f_1 = f_2 = \cdots = f_\ell = 0$ is a set of partially commutative polynomial equations from $\mathbb{F}\langle X \rangle$ and suppose that the following equations (axioms) are included among the f_i's.

- ***Boolean axioms**: $x^2 - x$ for every $x \in X$*
- ***Commutator axioms**: $x_{i,j}x_{i,j'} - x_{i,j'}x_{i,j}$ for every $i \in [m]$, $j \neq j' \in [n_i]$.*

A partially commutative IPS *proof (*$\mathsf{PC}_{m,n}$-IPS*) of unsatisfiability of the system $\{f_i\}$ is a polynomial $P(X, z_1, \ldots, z_\ell)$ represented by a partially commutative formula over the quotient of $\mathbb{F}\langle X, Z \rangle$ by the prescribed commuting relations among the X-variables only; in particular, no additional commuting relations involving the placeholder variables $Z = \{z_1, \ldots, z_\ell\}$ are imposed. We require*

1. *$P(X, \overline{0}) = 0$;*

[7] For every i, the variables within X_i are noncommuting, but for every $i \neq j$, variables from X_i and X_j commute.

2. $P(X, f_1, \ldots, f_\ell) = 1$,

where substitution is the algebra homomorphism that fixes the X-variables and sends $z_i \mapsto f_i$. The size *of a* $\mathsf{PC}_{m,n}$-IPS *proof is the minimum size of a partially commutative formula computing $P(X, Z)$.* ◊

Our Result Both commutative and noncommutative IPS are known to be sound and complete proof systems [GP18, LTW18]. Partially commutative IPS is an intermediate model; we give a formal proof of its soundness and completeness in Subsection A.1. Verification of noncommutative IPS proofs can be done in deterministic polynomial-time using the Raz–Shpilka noncommutative ABP identity-testing algorithm [RS05, LTW18]. On the other hand, commutative IPS has randomized polynomial-time verification via polynomial identity testing [Zip79]. For two buckets over $\mathbf{GF}(2)$, deterministic polynomial-time verification follows from Theorem 1.3; more generally, for a constant number of buckets over $\mathbb{Q}$, deterministic polynomial-time identity testing is known from [ACM24]. We refer the reader to Subsection A.2 for a more detailed discussion. We also show in Subsection A.3 that partially commutative IPS can efficiently simulate the Frege proof system, in particular the Łukasiewicz calculus described in Definition 2.3.

Our main result is a quasi-polynomial simulation of a natural subclass of partially commutative IPS by Frege. In particular, let $X = X_1 \sqcup X_2$ be a set of partially commutative variables with $|X_1| = n_1$, $|X_2| = n_2$, and $n_1 + n_2 = n$. Then, we show that any $\mathsf{PC}_{2,n}$-IPS proof of unsatisfiability for a CNF, can be converted into a Frege proof with at most a quasi-polynomial increase in size. This extends the simulation theorem of Li, Tzameret, and Wang from the fully noncommutative setting to a substantially more general, partially commutative model.

We fix $\mathbb{F}$ to be $\mathbf{GF}(2)$ throughout. For notational simplicity, we assume that $|X_1| = |X_2| = n'$. The arguments do not depend on the two buckets having the same size, and the same proofs apply to an arbitrary partition $X = X_1 \sqcup X_2$.

Theorem 1.2. *Any size-s* $\mathsf{PC}_{2,n}$-IPS *refutation over $\mathbb{F}$ of an unsatisfiable CNF can be simulated by a* Frege *proof of size $s^{(\log s)^{O(1)}}$.*

Note that the noncommuting graph in this setting is the disjoint union of two cliques of size n'. As discussed earlier, proving Theorem 1.2 reduces to establishing the soundness of $\mathsf{PC}_{2,n}$-IPS within Frege. Equivalently, we need efficient Frege proofs of the Boolean tautologies arising from partially commutative polynomial identities.

At first sight, one might hope to use an existing deterministic Polynomial Identity Testing algorithm for partially commutative polynomials. However, this is not sufficient for our purposes. To obtain short Frege proofs, one needs an identity testing algorithm whose correctness can itself be efficiently formalized within Frege. Existing algorithms rely on sophisticated algebraic constructions, and it is currently unclear how to efficiently formalize their correctness in Frege.

We therefore develop a new deterministic white-box identity testing algorithm for partially commutative ABPs. The algorithm is designed together with its proof of correctness, so that every

step admits an efficient Frege formalization. This forms the main technical ingredient in our proof of Theorem 1.2.

Theorem 1.3. (White-box algorithm for 2-bucket) *Let A be a partially commutative ABP in $\mathbb{F}\langle X_1 \sqcup X_2\rangle$ of size s and depth d. Then we can decide deterministically in time $\mathrm{poly}(s, d, n)$ whether A computes the zero polynomial. Moreover the correctness of the algorithm can be efficiently proved in Frege.*

Informal Description of the Algorithm: Our white-box algorithm is inspired by the noncommutative identity testing algorithm of Raz and Shpilka [RS05]. Recall that the variables are partitioned into two buckets X_1 and X_2, where variables within the same bucket do not commute, while variables from different buckets commute. Using Lemma B.17, we first transform the given partially commutative ABP into one that is homogeneous with respect to the variables in X_1. Consequently, every edge label has the form $\sum_{j=1}^{n'} x_{1,j} f_j$, where each $f_j \in \mathbb{F}\langle X_2\rangle$ is computed by a polynomial-size noncommutative ABP.

Let A be such a homogeneous partially commutative ABP of width w, depth d, source s, and sink t. For every layer $i \in [d]$, let $\mathbf{P}_i \in \mathbb{F}\langle X_1 \sqcup X_2\rangle^w$ denote the vector of partial ABPs, where the j-th coordinate is the sub-ABP from the source s to the j-th node in layer i. Thus every coordinate of $\mathbf{P}_i$ computes a polynomial of X_1-degree exactly i, and the vector $\mathbf{P}_i$ captures the computation performed by the ABP up to the i-th layer.

For every monomial $m \in \mathbb{F}\langle X_1\rangle$, define the coefficient vector $\mathsf{coeff}_m(\mathbf{P}_i) \in \mathbb{F}\langle X_2\rangle^w$, whose j-th coordinate is the coefficient of m in the polynomial computed by the j-th partial ABP. Equivalently, $\mathsf{coeff}_m(\mathbf{P}_i)$ records how the monomial m propagates simultaneously through every node of the i-th layer while preserving the remaining computation over the variables in X_2.

Let

$$\mathcal{V}_i = \{\mathsf{coeff}_m(\mathbf{P}_i) \ : \ m \in \mathbb{F}\langle X_1\rangle\} \subseteq \mathbb{F}\langle X_2\rangle^w$$

denote the collection of all coefficient vectors at layer i. Since there are exponentially many monomials in X_1, the set $\mathcal{V}_i$ is exponentially large. Therefore, the algorithm never stores $\mathcal{V}_i$ explicitly. Instead, it maintains a basis $\mathcal{B}_i \subseteq \mathbb{F}\langle X_2\rangle^w$ such that every coefficient vector in $\mathcal{V}_i$ can be written as a left $\mathbb{F}\langle X_2\rangle$-linear combination of vectors from $\mathcal{B}_i$. In addition, the vectors in $\mathcal{B}_i$ and $\mathcal{V}_i$ have the same left span over the free skew field $\mathbb{F}\langle\langle X_2\rangle\rangle$. Thus, $\mathcal{B}_i$ provides a succinct representation of all coefficient vectors at layer i.

The crucial observation is that these basis sets can be propagated recursively from one layer to the next. Indeed, if

$$\mathbf{P}_{i+1} = \mathbf{P}_i \left(\sum_{j=1}^{n'} x_{1,j} M_j^{(i)} \right),$$

where the matrices $M_j^{(i)}$ describe the transitions from layer i to layer $i+1$, then for every monomial $m \in \mathbb{F}\langle X_1\rangle$,

$$\mathsf{coeff}_{m x_{1,j}}(\mathbf{P}_{i+1}) = \mathsf{coeff}_m(\mathbf{P}_i) M_j^{(i)}.$$

Hence, every coefficient vector at layer $i+1$ is obtained by multiplying a coefficient vector at layer i with one of the transition matrices. Consequently, starting from a basis $\mathcal{B}_i$, we obtain a spanning set for $\mathcal{V}_{i+1}$ simply by multiplying every basis vector with each transition matrix. The remaining task is to compress this spanning set back into a basis. This is precisely where our main technical ingredient enters.

Lemma 1.4. (Factorization Lemma (Informal)) *Let $M \in \mathbb{F}\langle X\rangle^{\ell\times m}$ be a matrix of noncommutative polynomials of noncommutative rank r. Assume every entry of M is computed by a polynomial-size noncommutative ABP. Then one can efficiently compute a factorization $M = GH$, where $G \in \mathbb{F}\langle X\rangle^{\ell\times r}$, $H \in \mathbb{F}\langle X\rangle^{r\times m}$, and every entry of both G and H is computed by a polynomial-size noncommutative ABP. Moreover, H has full noncommutative row rank. Consequently, every row of M is a left $\mathbb{F}\langle X\rangle$-linear combination of the rows of H, and the rows of M and H have the same left span over the free skew field $\mathbb{F}\langle\langle X\rangle\rangle$.*

Remark 1.5. *The proofs of Theorem 1.3 as well as the factorization lemma (Lemma 1.4) continue to hold when $\mathbb{F}$ is a finite field or $\mathbb{Q}$.* ◊

The Factorization Lemma provides exactly the compression step required by the recursive algorithm. At every layer, we arrange the spanning set generated from $\mathcal{B}_i$ as the rows of a matrix over $\mathbb{F}\langle X_2\rangle$. The Factorization Lemma computes a rank factorization of this matrix, from which we obtain a full-row-rank spanning set while preserving succinct ABP representations and the same skew-field span. This basis becomes $\mathcal{B}_{i+1}$ and is propagated to the next layer. Repeating this procedure throughout the ABP yields a deterministic polynomial-time white-box identity testing algorithm for partially commutative ABPs over two buckets.

The algorithm is designed not only to test polynomial identities efficiently but also to allow for an efficient formalization within *Frege*. Both the layer-by-layer propagation and the factorization procedure can be verified by polynomial-size *Frege* proofs. This forms the first step towards proving the soundness of $\mathsf{PC}_{2,\mathrm{n}}$-$\mathsf{IPS}$ within *Frege*.

More precisely, let $P(X, Z)$ be a $\mathsf{PC}_{2,\mathrm{n}}$-$\mathsf{IPS}$ proof of an unsatisfiable CNF, where $Z = (z_1, \ldots, z_\ell)$ are the placeholder variables. By definition, P satisfies the identities

$$P(X, \bar{0}) = 0, \qquad P(X, f_1, \ldots, f_\ell) = 1,$$

where $f_1, \ldots, f_\ell$ are the polynomial encodings of the clauses, boolean axioms and commutator axioms. After Booleanization, these become the tautologies

$$\neg\widetilde{P}(X, \bar{0}), \qquad \widetilde{P}(X, \widetilde{f_1}, \ldots, \widetilde{f_\ell}).$$

The reflection principle states that polynomial-size *Frege* proofs of these two tautologies yield a *Frege* refutation of the original CNF.

Thus the remaining task is to prove the Booleanizations of partially commutative polynomial identities efficiently in *Frege*. For this, one needs suitable *identity witnesses* certifying that a partially

commutative ABP computes the zero polynomial, together with polynomial-size *Frege* proofs of their correctness. Our white-box identity testing algorithm and the Factorization Lemma provide exactly the machinery needed to construct such witnesses. The following theorem is the key technical result of this paper.

Lemma 1.6. (Witness Theorem for Partially Commutative ABP Identities) *Let F be a field and A be a partially commutative homogeneous (over X_1) ABP in $\mathbb{F}\langle X_1 \sqcup X_2\rangle$ of size s computing the identically zero polynomial. Let d be the depth of A, and let w_i denote the width of the i-th layer. Then there exist $2d-1$ matrices $\{\lambda'_0, \ldots, \lambda'_{d-1}\}$ and $\{T'_1, \ldots, T'_{d-1}\}$ such that*

$$\lambda'_i \in \mathbb{F}^{w_i \times w_i}\langle X_2\rangle, \qquad T'_i \in \mathbb{F}^{w_{i-1}\times w_i}\langle X_1 \sqcup X_2\rangle,$$

and the following hold:

$$\lambda'_{d-i} \cdot \overline{A}_i = \overline{0} \quad \forall\, i = 1, 2, \ldots, d, \tag{1.7}$$

$$\lambda'_{d-i} \cdot \overline{A}_i = T'_{d-i+1} \cdot \lambda'_{d-i+1} \cdot \overline{A}_{i-1} \quad \forall\, i = 2, 3, \ldots, d, \tag{1.8}$$

$$\lambda'_0 := 1, \qquad \lambda'_0 \cdot \overline{A}_d = \overline{A}_d = 0. \tag{1.9}$$

Moreover, every entry of the matrices λ'_i and T'_j is computable by a polynomial-size noncommutative ABP.

1.2 Proof Overview

The proof of Theorem 1.2 consists of two conceptual steps. The first, and technically most substantial, step is to construct identity witnesses for partially commutative ABPs over two buckets. These witnesses certify that a partially commutative ABP computes the zero polynomial and, moreover, admit efficient formalization within *Frege*. The second step is to use these witnesses to prove the soundness of $\mathsf{PC}_{2,\mathrm{n}}$-IPS inside *Frege*, which yields the desired quasi-polynomial simulation via the reflection principle. Our white-box identity testing algorithm is designed to admit efficient Frege formalization. The witness theorem can be viewed as the proof-complexity counterpart of this algorithm: it converts the recursive computation performed by the algorithm into algebraic certificates that can be verified efficiently inside Frege.

1.2.1 Proof Overview for the Construction of *Witness Identities*

The main technical contribution of the paper is the construction of witness identities for partially commutative ABPs over two buckets. We begin by recalling the witness construction of Li, Tzameret, and Wang for the noncommutative setting, since our proof can be viewed as a generalization of their argument.

Let A be a noncommutative ABP of depth d computing the identically zero polynomial, and let the width of layer i be w_i, with nodes $\{u_{i,1}, \ldots, u_{i,w_i}\}$. Also, let s, t be the source and sink nodes

in A. For $i \in [d]$, define $\overline{A}_i = \left[A[u_{d-i,1}, t], \ldots, A[u_{d-i,w_{d-i}}, t]\right]^\top$. Assuming A is homogeneous, each entry of $\overline{A}_i$ computes a homogeneous polynomial of degree i. We use $A[\ell_1, \ell_2]$ to denote the polynomials computed between the nodes ℓ_1, ℓ_2.

In this setting, [LTW18] constructs matrices $\{\lambda_1, \ldots, \lambda_d \mid \lambda_i \in \mathbb{F}^{w_i \times w_i}\}$ and $\{T_1, \ldots, T_d \mid T_i \in \mathbb{F}\langle X\rangle^{w_i \times w_{i+1}}\}$ witnessing the zeroness of the polynomial computed by A. They do so using an inductive argument based on Gaussian elimination over $\mathbb{F}$.

Formally, let M be the transition matrix from layer i to $i+1$. Note that it can be uniquely expressed as $M = \sum_{j=1}^{n} x_j M_j$. From the inductive hypothesis, $\lambda_i \cdot \overline{A}_{d-i} = 0$ and $\overline{A}_{d-i} = M \cdot \overline{A}_{d-i-1}$. So, $\sum_{j=1}^{n} x_j(\lambda_i M_j) \cdot \overline{A}_{d-i-1} = 0$, and hence, by homogeneity and noncommutativity, $(\lambda_i M_j) \cdot \overline{A}_{d-i-1} = 0$ for $j \in [n]$.

Let $V = \mathrm{Span}_{\mathbb{F}}\{\mathrm{Rows}(\lambda_i M_j) \mid j \in [n]\} \subseteq \mathbb{F}^{w_{i+1}}$. Every vector in V is orthogonal to $\overline{A}_{d-i-1}$. Choosing a basis matrix of V, say B, by Gaussian elimination on the stacked matrix $(\lambda_i M_1; \ldots; \lambda_i M_n)$, we obtain matrices T_j such that $\lambda_i M_j = T_j B$ for all $j \in [n]$. Now if we define $T_{i+1} := \sum_{j=1}^{n} x_j T_j$, the required properties are satisfied. Since all matrices λ_{i+1} and $T_1, \ldots, T_n$ have entries in $\mathbb{F}$, this is sufficient for [LTW18] regarding expressiveness in *Frege*.

We now explain why this argument does not extend directly to partially commutative ABPs. If we attempt to apply the same idea directly, the first important point is that the matrices λ_i and the coefficient matrices M_j are no longer scalar matrices over $\mathbb{F}$, but rather matrices whose entries are noncommutative polynomials over $\mathbb{F}\langle X_2\rangle$.

Thus, the rows of $\lambda_i M_j$ span a left module over the free skew field $\mathbb{F}\langle\langle X_2\rangle\rangle$. If one were to perform Gaussian elimination in this setting, the resulting matrices T_j would have entries in $\mathbb{F}\langle\langle X_2\rangle\rangle$, potentially involving nested inverses. Such expressions do not admit canonical representations as noncommutative polynomials[8], and therefore cannot be handled explicitly inside *Frege*. A standard example is $(z + xy^{-1}x)^{-1} - z^{-1}$ whcih cannot be written as $\frac{f}{g}$ where both $f, g \in \mathbb{F}\langle x, y, z\rangle$.

To overcome this obstacle, we construct the witness matrices directly within $\mathbb{F}\langle X_1 \sqcup X_2\rangle$ using a factorization-based inductive procedure. Conceptually, this factorization replaces Gaussian elimination over the *free skew field* by an explicit decomposition within $\mathbb{F}\langle X_1 \sqcup X_2\rangle$. Instead of computing a basis via division, which would introduce inverses, we obtain a structured factorization that preserves polynomial representations and allows the construction to be simulated within Frege.

The proof proceeds by induction on $r = 0, 1, \ldots, d-1$, constructing matrices λ'_r and T'_{r+1} such that

$$\lambda'_{d-i} \cdot \overline{A}_i = \overline{0} \quad \forall\, i = 1, 2, \ldots, d,$$

and

$$\lambda'_{d-i} \cdot \overline{A}_i = T'_{d-i+1} \cdot \lambda'_{d-i+1} \cdot \overline{A}_{i-1} \quad \forall\, i = 2, 3, \ldots, d,$$

[8]The entries from $\mathbb{F}\langle\langle X\rangle\rangle$ do not even admit a canonical representation like commutative rational functions.

where $\lambda'_i \in \mathbb{F}\langle X_2\rangle^{w_i \times w_i}$ and $T'_i \in \mathbb{F}\langle X_1 \sqcup X_2\rangle^{w_{i-1} \times w_i}$.

The matrices λ'_r can be viewed as encoding linear dependencies among the coefficient vectors at layer $d-r$. In particular, the invariant $\lambda'_r \cdot \overline{A}_{d-r} = 0$ expresses that the rows of λ'_r lie in the space of relations among the polynomials computed at that layer. The base case $r = 0$ follows from $\overline{A}_d = 0$. The matrices λ'_r are chosen by the factorization procedure so that the transition identities can be propagated across layers; they need not be nonzero in every degenerate rank case.

For the inductive step, assume $\lambda'_r \cdot \overline{A}_{d-r} = 0$. Using homogeneity over X_1, we obtain

$$(\lambda'_r M_j) \cdot \overline{A}_{d-r-1} = 0 \quad \forall j \in [n].$$

Define the stacked matrix

$$T = \begin{pmatrix} \lambda'_r M_1 \\ \vdots \\ \lambda'_r M_n \end{pmatrix}.$$

Then $\overline{A}_{d-r-1}$ lies in the right kernel of T.

Instead of performing Gaussian elimination over the free skew field, we compute a *constructive matrix factorization*

$$T = C \cdot B$$

using the factorization procedure of Lemma 3.5.

This factorization is obtained by first linearizing T via a variant of Higman's construction (see Lemma 3.15), and then applying a structured decomposition of the resulting linear matrix. The procedure yields matrices C and B with polynomial-sized ABP entries.

Moreover, the factorization satisfies that for each $j \in [n]$, the corresponding block C_j of C satisfies

$$\lambda'_r M_j = C_j B.$$

We now define

$$\lambda'_{r+1} := B, \qquad T'_{r+1} := \sum_{j=1}^{n} x_{1,j} C_j.$$

This yields the transition identity

$$\lambda'_r \cdot \overline{A}_{d-r} = T'_{r+1} \cdot \lambda'_{r+1} \cdot \overline{A}_{d-r-1}.$$

To obtain the next vanishing condition, we use the rank properties of the factorization. If C has full column rank, then

$$C \cdot B \cdot \overline{A}_{d-r-1} = 0 \;\Rightarrow\; \lambda'_{r+1} \cdot \overline{A}_{d-r-1} = 0.$$

Otherwise, the structure of the decomposition forces $\overline{A}_{d-r-1} = 0$, implying that $\lambda'_{r+1} \cdot \overline{A}_{d-r-1} = 0$.

This completes the inductive construction. Finally, since all matrices involved are computed by

polynomial-size ABP, and the factorization preserves this complexity, every entry of the witness matrices λ_i' and T_j' admits a polynomial-size ABP. The witness identities certify the correctness of identity testing for partially commutative ABPs in a form that can be efficiently verified inside Frege. We now explain how these witnesses are used to prove the soundness of $\mathsf{PC}_{2,n}$-IPS.

1.2.2 Proof Overview for the Frege Simulation

As mentioned earlier, the field $\mathbb{F}$ is fixed to be $\mathbf{GF}(2)$ and we assume the witness identities of Lemma 1.6. Our goal is to prove Theorem 1.2, namely that every size-s $\mathsf{PC}_{2,n}$-IPS refutation of an unsatisfiable CNF

$$\Phi = \kappa_1 \wedge \kappa_2 \wedge \cdots \wedge \kappa_\ell$$

can be simulated by a quasi-polynomial-size *Frege* refutation.

Using the standard arithmetization, we obtain an unsatisfiable system of partially commutative polynomials

$$P_\Phi = \{P_{\Phi,1}, \ldots, P_{\Phi,\ell}\}.$$

Let $F(X, Z)$ be a $\mathsf{PC}_{2,n}$-IPS refutation of this system, where $X = X_1 \sqcup X_2$ and Z denotes the placeholder variables. Let $\overline{P_\Phi}$ be the set of polynomials obtained by adding the boolean axioms and commutator axioms to P_Φ. As mentioned earlier, we assume $|X_1| = |X_2| = n'$, By definition,

$$F(X, \overline{0}) = 0, \qquad F(X, \overline{P_\Phi}) = 1.$$

Applying the Booleanization map (Definition 3.1) yields Boolean tautologies.

$$\neg\widetilde{F}(X, \overline{0}), \qquad \widetilde{F}(X, \widetilde{P_\Phi}).$$

By the reflection principle (Lemma B.1), it is sufficient to provide quasi-polynomial-size Frege proofs of these two tautologies.

Consequently, it is enough to prove that whenever a partially commutative formula computes the zero polynomial, its Booleanization has a quasi-polynomial size Frege proof. This is precisely the content of Theorem B.2.

The proof begins by applying homogenization using Lemma B.3 to the given partially commutative formula, followed by Nisan's construction to obtain a layered homogeneous partially commutative ABP A_F. We then partially homogenize A_F with respect to the bucket X_1 using Lemma B.17. For every $r \geq 1$, this gives a compressed layered ABP $A'_{F,r}$ whose source-to-sink polynomial is the X_1-homogeneous component of degree r; the degree-zero component is a noncommutative polynomial over X_2. Every edge label of $A'_{F,r}$ is a homogeneous linear form in X_1 whose coefficients are polynomial-size noncommutative ABPs over X_2.

The main technical step is to associate every node of each $A'_{F,r}$ with an induced formula derived

from the original homogeneous formula. Starting from the induced formulas of Li, Tzameret and Wang [LTW18] for the homogeneous ABP A_F, we refine the construction so that every duplicated node of $A'_{F,r}$ is tracked by an induced formula computing the same polynomial. Furthermore, these induced formulas satisfy local ABP identities corresponding to the recursive computation of $A'_{F,r}$, and these identities admit polynomial-size Frege proofs.

For every $r \geq 1$, the compressed ABP $A'_{F,r}[s, t^{(r)}]$ computes the X_1-homogeneous component of degree r. Since the original formula computes the zero polynomial, each of these ABPs also computes the zero polynomial, and we apply the witness theorem (Lemma 1.6) to obtain witness matrices satisfying the vanishing and transition identities. The degree-zero component is a noncommutative identity over X_2 and is handled directly using [LTW18].

The witness identities are Booleanized and proved inside Frege. The proof reduces these identities to polynomially many noncommutative ABP identities over X_2. These noncommutative identities are then proved using the Frege simulation of noncommutative identity testing due to Li, Tzameret and Wang [LTW18].

Finally, the Booleanized witness identities are propagated layer by layer through the partially homogeneous ABP. The vanishing identities provide the base case, while the transition identities propagate the argument from one layer to the next, yielding a Frege proof of $\neg\widetilde{F^*_{s,r}}$ for every homogeneous component. Using the homogeneous decomposition of F, these derivations are combined to obtain a Frege proof of $\neg\widetilde{F}$, and the reflection principle completes the simulation of the original $\mathsf{PC}_{2,n}$-IPS refutation.

The resulting Frege proof has quasi-polynomial-size. This follows because the partially homogeneous ABP has polynomial size, every induced formula has size $s^{O(\log^2 s)}$, every witness entry is computable by a polynomial-size noncommutative ABP, and the layer-by-layer propagation introduces only a polynomial overhead.

1.3 Further Remarks

The main contribution of this paper is a quasi-polynomial simulation of $\mathsf{PC}_{2,n}$-IPS by the Frege proof system. In this sense, we make some progress towards understanding how much commutativity Frege can handle while simulating IPS.

A natural question is whether the simulation extends to $\mathsf{PC}_{k,n}$-IPS for constant k. Arvind, Chatterjee, and Mukhopadhyay [ACM24] give a deterministic polynomial-time identity-testing algorithm for partially commutative ABPs over $X_1 \sqcup \cdots \sqcup X_k$, for constant k and over $\mathbb{Q}$. Their algorithm proceeds through the more general singularity-testing problem for linear matrices over the corresponding partially commutative algebra. It is not clear how to obtain from this algorithm identity witnesses of the form required for an efficient Frege formalization over $\mathbf{GF}(2)$. A further obstacle is the construction of skew fields containing partially commutative rings. The algorithm of [ACM24] relies on a skew-field construction over $\mathbb{Q}\langle X_1 \sqcup \cdots \sqcup X_k\rangle$, which is presently

known only over fields of characteristic zero when $k \geq 3$ [KVV20]. Moreover, the factorization based technique to formulate the efficient witness from this work seems not extendable to three or more buckets. Formally speaking, for two buckets, after homogenizing with respect to X_1, the coefficients lie in the free algebra $\mathbf{GF}(2)\langle X_2 \rangle$. Our witness construction uses rank factorizations of matrices over this free algebra given by Cohn. Such efficiently representable factorization holds since the noncommutative ring is a *free ideal ring* [Coh06, Chapter 2].

For three buckets, the same approach would instead require working over the coefficient ring $\mathbf{GF}(2)\langle X_1, X_2 \rangle$ and, more generally, over a partially commutative algebra when the number of buckets is larger. Universal skew fields of fractions for these algebras are known to exist [KVV20]. However, partially commutative algebras with at least two unbounded sized partitions are not free ideal rings [KVV20]. Therefore, the rank-factorization argument used in Subsection 3.4 does not directly extend to these coefficient rings.

In particular, extending our proof to more buckets would require either an analogue of the matrix-factorization witnesses used here for multipartite coefficient rings, or a different form of identity witness that can be efficiently verified in Frege. We leave the following question open.

For every constant k, can $\mathsf{PC}_{k,n}$-IPS be quasi-polynomially simulated by Frege over $\mathbf{GF}(2)$?

Organization.

The remainder of the paper is organized as follows. Section 2 introduces the computational models and proof systems used in the paper. Section 3 develops the new white-box identity-testing algorithm, proves the Factorization Lemma, and constructs the witness matrices. Appendix A defines the partially commutative Ideal Proof System, establishes soundness and completeness, and explains how it simulates Frege. Finally, Appendix B gives the detailed Frege simulation and completes the proof of the main theorem.

2 Preliminaries

Notation. For any $i, j \in \mathbb{N}$ with $i \leq j$, we denote the set $\{i, \ldots, j\}$ by $[i, j]$. In addition, for any $n \in \mathbb{N}$, we denote the set $\{1, \ldots, n\}$ by $[n]$.

2.1 Computational Models

Definition 2.1 (Algebraic Circuits and Formulas)**.** *An algebraic circuit C is a directed acyclic graph with a unique output gate (root) of out-degree 0, and input gates of in-degree 0 (leaves) labeled by variables $x_1, \ldots, x_n$ or constants from $\mathbb{F}$. The internal gates are labeled by $+$ or $\times$.*

Each gate v computes a polynomial f_v defined recursively: if v is an input, then $f_v = \mathsf{label}(v) \in \{x_1, \ldots, x_n\} \cup \mathbb{F}$; if $v = u \ \mathsf{op} \ w$ for $\mathsf{op} \in \{+, \times\}$, then $f_v = f_u \ \mathsf{op} \ f_w$. The polynomial computed at the

output gate is the polynomial computed by the circuit.

If the underlying graph is restricted to be a tree, then the model is that of an algebraic formula. The size of an algebraic circuit or formula is the number of wires in it. ◊

Definition 2.2 (Algebraic Branching Program)**.** *An* algebraic branching program *(ABP) is a layered directed acyclic graph. The vertex set is partitioned into layers* $0, 1, \ldots, L$*, with directed edges only between adjacent layers (*i *to* $i+1$*). There is a* source *vertex of in-degree* 0 *in layer* 0*, and one out-degree-*0 sink *vertex in layer* L*. Each edge is labeled by an affine* $\mathbb{F}$*-linear form where* $\mathbb{F}$ *is the underlying field. The polynomial computed by the ABP is the sum over all source-to-sink directed paths of the ordered product of affine linear forms labeling the path edges.*

The size of an ABP is the number of vertices in it. ◊

Partially Commutative Models. Given a partition $X = X_1 \sqcup \cdots \sqcup X_m$ of the variables, an ABP is said to be Partially Commutative with respect to this partition if the multiplication respects the partial commutativity defined by the partition during its computation. Partially Commutative Formulas with respect to a given partition is defined analogously. Clearly, any polynomial computed by such an ABP or formula lies in $\mathbb{F}\langle X_1 \sqcup \cdots \sqcup X_m \rangle$.

Let A be a partially commutative ABP with respect to the partition $X_1 \sqcup X_2$ computing a polynomial $\hat{A} \in \mathbb{F}\langle X_1 \sqcup X_2 \rangle$. We will consider the edge labels in compressed form obtained after homogenizing with respect to the bucket X_1: each edge label is written as $\sum_{j=1}^{n} x_{1,j} \cdot f_j$, where each coefficient $f_j \in \mathbb{F}\langle X_2 \rangle$ is represented by a polynomial-size noncommutative ABP.

Further, we will denote the number of layers in A by d and the width of A by w. Thus $\deg_{X_1}(\hat{A}) = d$. Finally, we will denote the nodes in layer i by $\{u_{i,1}, u_{i,2}, \ldots, u_{i,w_i}\}$ and the source, sink nodes respectively by s, t. We can then define the following vector of ABPs:

$$\overline{A}_i := [A[u_{d-i,1}, t], A[u_{d-i,2}, t], \cdots, A[u_{d-i,w_{d-i}}, t]]^\top$$

where $A[u_{d-i,j}, t]$ is the partial ABP between $u_{d-i,j}$ and t for every $j \in [w_{d-i}]$.

Observe that each ABP in $\overline{A}_i$ computes a partially commutative polynomial which is homogeneous over X_1 and has X_1-degree i.

2.2 The Ideal Proof Systems

We start with the definition of the Ideal proof system (IPS) as defined in [GP18].

Definition 2.3 (Ideal Proof Systems [FSTW21, GP18, LTW18])**.** *Let* $f_1, \ldots, f_\ell \in \mathbb{F}[X]$ *be a set of polynomials such that* $\{f_1, \ldots, f_\ell, x_1^2 - x_1, \ldots, x_n^2 - x_n\}$ *has no common solution*[9]*. A* proof *of the unsatisfiability of this set of polynomial equations, in the* Ideal Proof System (IPS)*, is a polynomial* $P(X, y_1, \ldots, y_\ell, z_1, \ldots, z_n) \in \mathbb{F}[X, Y, Z]$ *such that the following holds:*

[9]That is, there does not exist $\bar{x} \in \{0,1\}^n$ such that for every $i \in [\ell]$, $f_i(\bar{x}) = 0$.

- $P(X, \overline{0}, \overline{0}) = 0$;
- $P(X, f_1, \ldots, f_\ell, x_1^2 - x_1, \ldots, x_n^2 - x_n) = 1$.

The size *of an* IPS *proof is the minimal size of a circuit (formula or ABP) computing* $P(X, Y, Z)$.

If $f_1, \ldots, f_\ell \in \mathbb{F}\langle X \rangle$ *are, instead, a set of noncommutative polynomials such that* $\{f_1, \ldots, f_\ell\} \cup \{x_i^2 - x_i : i \in [n]\} \cup \{x_i x_j - x_j x_i : i, j \in [n]\}$ *has no common solutions, then a* proof *of the unsatisfiability of this set of polynomial equations, in the* noncommutative Ideal Proof System (NC-IPS)*, is a polynomial* $P(X, y_1, \ldots, y_\ell, z_1, \ldots, z_n, w_1, \ldots w_{n^2}) \in \mathbb{F}\langle X, Y, Z, W \rangle$ *such that the following holds:*

- $P(X, \overline{0}, \overline{0}) = 0$;
- $P(X, f_1, \ldots, f_\ell, x_1^2 - x_1, \ldots, x_n^2 - x_n, x_1x_2 - x_2x_1, \ldots, x_{n-1}x_n - x_nx_{n-1}) = 1$.

The size *of an* NC-IPS *proof is the minimal size of a noncommutative formula computing* $P(X, Y, Z, W)$. ♢

Remark 2.4. *Suppose we are given a system of equations, say* $\{f_1 = 0, \ldots, f_\ell = 0 : f_i \in \mathbb{F}[x_1, \ldots, x_n]\}$*, in the commutative setting. We say that the system is unsatisfiable if there is no* $\overline{a} \in \mathbb{F}^n$ *such that* $f_i(\overline{a}) = 0$ *for every* $i \in [\ell]$*. However, when we consider the variable set* X *to be noncommuting, we need to look for a common solution in the matrix algebra over* $\mathbb{F}$*, denoted by* $\mathbb{F}^{d \times d}$ *for some* $d > 1$.

In particular, elements from $\mathbb{F}$ *are thought of as* 1×1 *dimensional matrices. Hence, given a set of noncommutative equations* $\{f_1 = 0, \ldots, f_m = 0 : f_i \in \mathbb{F}\langle x_1, \ldots, x_n \rangle\}$*, we say that they are unsatisfiable if there does not exist any* d*,* $\{\overline{A} = (A_1, \ldots, A_n) : A_i \in \mathbb{F}^{d \times d}\}$ *such that* $f_i(\overline{A}) = \mathbf{0}$ *for every* $i \in [m]$. ♢

Example 2.5 (Evaluation in Matrix algebra)**.** $xy - yx + 2$ *evaluates to*

$$\begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} \cdot \begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix} - \begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix} \cdot \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} + 2I_{2\times 2}$$

when we substitute $x = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix}, y = \begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix} \in \mathbb{F}^{2\times 2}$. ♢

2.3 The Frege Proof System

Definition 2.6 (Boolean Formula)**.** *A* Boolean formula Φ *over variables* $x_1, \ldots, x_n$ *is a finite rooted tree with a fan-in of at most* 2 *in which:*

- *Each leaf is labeled by either a variable* x_i *or a Boolean constant* $\{0, 1\}$.
- *Each internal node is labeled by one of the Boolean connectives* $\{\vee, \wedge, \neg\}$.
- *A node labeled by* $\neg$ *has exactly one child (unary connective).*
- *Nodes labeled by* $\vee$ *or* $\wedge$ *have exactly two children (binary connectives).*

The formula computes the Boolean function obtained by evaluating the tree in the natural way: each gate applies its connective to the values computed by its child/ children. The size of the formula is the number of nodes in it, denoted as $|\Phi|$. ◊

We now define the *Frege* proof system.

Definition 2.7. (*Frege* derivation rule, [LTW18, Definition 2.2], [CR79]) *Let* $\overline{x}$ *be a set of boolean variables. A* Frege *rule is a sequence of propositional formulas* $A_0(\overline{x}), \ldots, A_k(\overline{x})$*, for* $k \geq 0$*, written as* $\frac{A_1(\overline{x}),\ldots,A_k(\overline{x})}{A_0(\overline{x})}$*. When* $k = 0$*, the Frege rule is called axiom scheme.*

A formula F_0 *is said to be derived by the rules from* $F_1, \ldots, F_k$ *if there are formulas* $B_1, \ldots, B_n$ *such that for every* $i \in \{0, 1, , \ldots, k\}$*,* $A_i(B_1/x_1, \ldots, B_n/x_n) = F_i$[10]*. The rule is said to be sound if any assignment satisfying* $A_1(\overline{x}), \ldots, A_k(\overline{x})$ *also satisfies* $A_0(\overline{x})$. ◊

A proof system is sound if it admits proofs only of tautologies. It is said to be *implicationally complete* if, for all sets of formulas T, if T semantically imply F, then there is a derivation of F in the proof system from the axioms T.

Definition 2.8. (Frege Proof, [LTW18, Definition 2.3], [CR79]) *Given a set of Frege rules, a Frege proof of a boolean formula* A *is a sequence of boolean formulas such that every formula is either an axiom or derived by one of the rules from previous formulas and terminates at* A*. The size of the proof is the sum of the sizes of all formulas in the proof. Given a set of sound Frege rules* P*, we say* P *is a Frege proof system if it is* implicationally complete. ◊

It is known from the work of Reckhow [Rec76], that all *Frege* proof systems are polynomially equivalent to each other. We now describe the instantiation of the *Frege* proof system that we will work with in this paper, known as the Łukasiewicz calculus. For propositional formulas A, B, C, we abbreviate $\neg A \vee B$ as $A \to B$.

The system has only three axiom schemes, and they are as follows.

1. $A \to (B \to A)$.
2. $(\neg A \to \neg B) \to ((\neg A \to B) \to A)$.
3. $(A \to (B \to C)) \to ((A \to B) \to (A \to C))$.

The single inference rule (known as *modus ponens*) used here is $\frac{A, A \to B}{B}$. That is, from A and $A \to B$ we infer B.

3 Quasi-Polynomial Simulation of $\mathsf{PC}_{2,\mathrm{N}}$-$\mathsf{IPS}$ via Frege using a new PIT Algorithm for Partially Commutative ABPs Over Two Buckets

In this section, we discuss our main technical contribution.

[10] we define $A_i(B_1/x_1, \ldots, B_n/x_n)$ by the formula A_i where x_i is substituted by B_i.

We begin by giving a more formal version of our main theorem (Theorem 1.2). As discussed in the proof overview Subsection 1.2, the proof of the Frege simulation closely resembles that of [LTW18], with each of the steps having to be carefully modified in order to get equivalent statements for the partially commutative setting. The most non-trivial modification required is that of the Witness Lemma (Lemma 1.6). Even though the statement looks very similar to the one in [LTW18], an entirely new identity testing algorithm for partially commutative ABPs with two buckets had to be developed in order for us to prove the existence of these witness identities. Therefore, we focus on describing this new algorithm here along with its proof of correctness.

A complete description of the quasi-polynomially large Frege simulation of $\mathsf{PC}_{2,\mathrm{N}}$-$\mathsf{IPS}$ (that is, the proof of Theorem 1.2) using these witness identities can be found in Appendix B.

3.1 Formal Statement of our Main Theorem

Given a boolean formula Ψ over boolean variables X with $|X| = N$ with internal gates $\wedge, \vee, \neg$, we define the corresponding partially commutative formula $\mathrm{Tr}(\Psi)$ in the following inductive way:

1. Partition the variables into $X = X_1 \sqcup X_2$ where $|X_1| = n_1$, $|X_2| = n_2$ and $n_1 + n_2 = N$. We treat X as a partially commutative algebraic variables now where variables from X_1 commute with variables from X_2 and vice versa, but variables within X_1 (similarly, X_2) are noncommuting.
2. $\mathrm{Tr}(x_i) \coloneqq 1 - x_i$ and $\mathrm{Tr}(\neg x_i) \coloneqq x_i$ for every $x_i \in X$.
3. $\mathrm{Tr}(\mathrm{False}) \coloneqq 0, \mathrm{Tr}(\mathrm{True}) \coloneqq 1$, $\mathrm{Tr}(\psi_1 \vee \psi_2 \vee \ldots \vee \psi_t) \coloneqq \prod_i \mathrm{Tr}(\psi_i)$. The product maintains the partial commutativity of the variables across the sets X_1, X_2.
4. Given a CNF $\Phi = \kappa_1 \wedge \kappa_2 \wedge \ldots \wedge \kappa_m$, we denote the set of partially commutative polynomials $\{\mathrm{Tr}(\kappa_1), \ldots, \mathrm{Tr}(\kappa_m)\}$ by P_Φ with $\mathrm{Tr}(\kappa_j) \coloneqq P_{\Phi,j}$.

Let F be a $\mathsf{PC}_{2,\mathrm{N}}$-$\mathsf{IPS}$ proof over $\mathbb{F}$ that witnesses the unsatisfiability of $P_\Phi = \{P_{\Phi,1}, \ldots, P_{\Phi,m}\}$, and let $\overline{P_\Phi}$ be the set obtained by adjoining the *boolean and commutator* axioms to P_Φ. Then, by Definition 1.1, F satisfies the conditions: $F(X, 0) = 0$ and $F(X, \overline{P_\Phi}) = 1$.

Definition 3.1. *Let $X = X_1 \sqcup X_2$ be partially commutative variables. Let F be a partially commutative formula computing a polynomial in $F\langle X\rangle$. We define the boolean formula $\widetilde{F}$ from F by changing the $+$ gate to $\oplus$, the $\times$ gate to $\wedge$, and treating the algebraic variables X as boolean variables.* $\Diamond$

Observe that over the boolean cube $\{0,1\}^N$, both formulas F and $\widetilde{F}$ evaluate to the same value. Now define the set $\widetilde{P}_\Phi \coloneqq \left\{\widetilde{P}_{\Phi,1}, \ldots, \widetilde{P}_{\Phi,m}, \widetilde{B}, \widetilde{C}\right\}$, where $\widetilde{B}$ and $\widetilde{C}$ is the set of boolean formulas we obtain from boolean axioms and commutator axioms over $\mathbb{F} = \mathbf{GF}(2)$. From the $\mathsf{PC}_{2,\mathrm{N}}$-$\mathsf{IPS}$ proof F, we obtain the following two tautologies using Definition 3.1:

$$\neg\widetilde{F}(X, 0) \qquad \widetilde{F}(X, \widetilde{P}_\Phi) \tag{3.2}$$

The formal statement of Theorem 1.2 is then the following.

Theorem 3.3. *Let $\Phi = \kappa_1 \wedge \cdots \wedge \kappa_m$ be an unsatisfiable CNF and $P_\Phi = \{\mathrm{Tr}(\kappa_1), \ldots, \mathrm{Tr}(\kappa_m)\}$ be the corresponding partially commutative system that has no common boolean root. If there is a $\mathsf{PC}_{2,\mathrm{N}}$-$\mathsf{IPS}$ refutation for $\overline{P_\Phi}$ of size s over $\mathbf{GF}(2)$ then there is a Frege proof for $\neg\Phi$ of size $s^{(\log s)^{O(1)}}$.*

As mentioned earlier, the proof is given in Appendix B. We now move on to the description our new identity testing algorithm for partially commutative ABPs with two buckets along with its proof of correctness. This would complete the proof of our second main theorem.

Theorem 1.3. (White-box algorithm for 2-bucket) *Let A be a partially commutative ABP in $\mathbb{F}\langle X_1 \sqcup X_2\rangle$ of size s and depth d. Then we can decide deterministically in time $\mathrm{poly}(s, d, n)$ whether A computes the zero polynomial. Moreover the correctness of the algorithm can be efficiently proved in Frege.*

The rest of this section is devoted to describing the algorithm and proving its correctness. The efficient Frege provability is given in Subsection B.6.

3.2 The White-Box Identity Testing Algorithm

We first describe the new deterministic white-box algorithm for testing whether a partially commutative ABP over two buckets computes the identically zero polynomial. Although $\mathbb{F}$ is fixed to be $\mathbf{GF}(2)$, we remark that similar statements continue to hold when $\mathbb{F}$ is a finite field or $\mathbb{Q}$.

Notational convention. For the remainder of this section, and only for notational simplicity, we assume that $|X_1| = |X_2| = n$. None of the arguments below use the fact that the two buckets have the same cardinality. For arbitrary $|X_1| = n_1$ and $|X_2| = n_2$, the same proofs go through verbatim by indexing the variables in X_1 by $[n_1]$ and those in X_2 by $[n_2]$. Thus the equal-size assumption has no effect on the statement or complexity of our results.

Setup. Let A be a layered partially commutative ABP over $\mathbb{F}\langle X_1 \sqcup X_2\rangle$ of size s and depth d, with source s and sink t. Carefully modifying the proof of a similar statement from the work of Li, Tzameret, Wang [LTW18], we can convert the ABP to another partially commutative ABP that is homogeneous over X_1 with a polynomial blow up in size (see Lemma B.17). Hence, we assume that A is homogeneous with respect to the X_1-variables and that every edge label is of the form

$$\sum_{j=1}^{n} x_{1,j} \cdot f_j,$$

where each f_j is computed by a noncommutative ABP over X_2 of size $\mathrm{poly}(s)$. The source s and sink t are in layer 0 and d, respectively. Let w_i be the number of nodes at layer i and assume

$w = \max_i \{w_i\}$. For the sake of simplicity, we assume each layer has w nodes (by adding dummy vertices wherever required).

For each layer $i \in \{0, \ldots, d\}$, define the vector of partial ABPs as

$$\mathbf{P}_i = \left(A[s, u_{i,1}], \ldots, A[s, u_{i,w}]\right) \in \mathbb{F}\langle X_1 \sqcup X_2 \rangle^w,$$

so that each entry of $\mathbf{P}_i$ is homogeneous of X_1-degree i.

Coefficient vectors. For a monomial $m \in \mathbb{F}\langle X_1 \rangle$, define

$$\mathsf{coeff}_m(\mathbf{P}_i) := \left(\mathsf{coeff}_m(A[s, u_{i,1}]), \ldots, \mathsf{coeff}_m(A[s, u_{i,w}])\right),$$

where each entry is a noncommutative polynomial over X_2. Let $\mathcal{V}_i$ denote the set of all such coefficient vectors, and $\mathbb{F}' = \mathbb{F}\langle\langle X_2 \rangle\rangle$ be the free skew field containing the ring $\mathbb{F}\langle X_2 \rangle$. The idea is to maintain a basis set $\mathcal{B}_i \subseteq \mathbb{F}\langle X_2 \rangle^w$ for every layer i such that

$$\text{every } u \in \mathcal{V}_i \text{ satisfies} \quad u = \sum_{v_j \in \mathcal{B}_i} p_j \cdot v_j \quad \text{with } p_j \in \mathbb{F}\langle X_2 \rangle. \tag{3.4}$$

In addition to Equation 3.4, we maintain the following linear-algebraic invariant over the free skew field $\mathbb{F}'$: the left $\mathbb{F}'$-linear span of $\mathcal{B}_i$ is exactly the left $\mathbb{F}'$-linear span of $\mathcal{V}_i$. That is, $\mathrm{span}_{\mathbb{F}'}(\mathcal{B}_i) = \mathrm{span}_{\mathbb{F}'}(\mathcal{V}_i)$. The stronger polynomial spanning property in Equation 3.4 will be used to propagate coefficient vectors from one layer to the next, whereas equality of the $\mathbb{F}'$-linear spans will be used in the final zero test.

We also need the following lemma to efficiently construct the factors of a polynomial matrix.

Lemma 3.5. (Factorization Lemma) *Let $M \in \mathbb{F}\langle X \rangle^{\ell \times m}$ be a matrix of noncommutative polynomials with $\mathrm{ncrank}(M) = r < \min\{\ell, m\}$. Assume each entry of M is computed by a size s noncommutative ABP. Then we can compute a nontrivial factorization of M, namely $M = G \cdot H$, such that the entries of $G \in \mathbb{F}\langle X \rangle^{\ell \times r}$ and $H \in \mathbb{F}\langle X \rangle^{r \times m}$ have $\mathrm{poly}(s, \ell, m)$-size noncommutative ABPs.*

We prove Lemma 3.5 in Subsection 3.3, but for now, we will use it to design the algorithm and prove its correctness.

Rank factorization convention. In the algorithm below, whenever we factor a polynomial matrix $C \in \mathbb{F}\langle X_2 \rangle^{a \times b}$ of noncommutative rank r, we use Lemma 3.5 when $r < \min\{a, b\}$. If $r = a \leq b$, we use the trivial factorization $C = I_a \cdot C$, and if $r = b \leq a$, we use $C = C \cdot I_b$. If $r = 0$, then $C = 0$ and the corresponding basis set is empty. Thus, in every case, we obtain a factorization $C = U \cdot V$ with inner dimension r, where U has full noncommutative column rank and V has full noncommutative row rank.

Algorithm (White-box PIT for Partially Commutative ABPs over Two Buckets).

Input: A layered, partially commutative ABP A as described above.
Output: ZERO if $A[s,t] \equiv 0$; otherwise, NONZERO.

1. **Initialization.** Consider the first layer. The set of vectors corresponding to layer 1, $\mathcal{V}_1 := \left\{ \mathsf{coeff}_{x_{1,j}}(\mathbf{P}_1) \ : \ j \in [n] \right\} \subseteq \mathbb{F}\langle X_2 \rangle^w$. Each entry of these vectors is computable by a noncommutative ABP of size $\mathrm{poly}(s)$. Consider the matrix M_1 of dimension $n \times w$ whose rows are the vectors from $\mathcal{V}_1$. Using the rank-factorization convention above, we obtain a factorization of the matrix $M_1 = C_1 \cdot B_1$ where C_1 and B_1 are noncommutative polynomial matrices of dimension $n \times r_1$ and $r_1 \times w$ with $r_1 = \mathrm{ncrank}(M_1)$. Moreover, the rows of B_1 are linearly independent over the field $\mathbb{F}'$. So we obtain $\mathcal{B}_1 := \mathrm{Rows}(B_1)$.

2. **For each layer** $i = 1, \ldots, d-1$**:**

   (a) Consider the matrices $M_j^{(i)}$ such that

   $$\mathbf{P}_{i+1} = \mathbf{P}_i \cdot \left(\sum_{j=1}^{n} x_{1,j} M_j^{(i)} \right).$$

   (b) For each $v \in \mathcal{B}_i$ and each $j \in [n]$, compute

   $$v^{(j)} := v \cdot M_j^{(i)}.$$

   (c) Form the matrix C_{i+1} whose rows are all vectors $v^{(j)}$.

   (d) Compute a factorization

   $$C_{i+1} = U \cdot V$$

   such that U has full column rank and V has full row rank over the free skew field. Every row of C_{i+1} is a left $\mathbb{F}\langle X_2 \rangle$-linear combination of the rows of V, and the two row sets have the same left span over $\mathbb{F}'$. Define $\mathcal{B}_{i+1} := \mathrm{Rows}(V)$.

3. **Final check.** If every vector in $\mathcal{B}_d$ has a zero entry in the coordinate corresponding to the sink node t, output ZERO; otherwise, output NONZERO.

Correctness. We show by induction on the layer i that the two invariants stated above hold:

(i) every vector in $\mathcal{V}_i$ is a left $\mathbb{F}\langle X_2 \rangle$-linear combination of vectors in $\mathcal{B}_i$,

(ii) $\mathrm{span}_{\mathbb{F}'}(\mathcal{B}_i) = \mathrm{span}_{\mathbb{F}'}(\mathcal{V}_i)$.

For $i = 1$, the factorization $M_1 = C_1 B_1$ gives the polynomial spanning property. Since B_1 has $r_1 = \mathrm{ncrank}(M_1)$ rows and full noncommutative row rank, its rows and the rows of M_1 have the same r_1-dimensional left span over $\mathbb{F}'$.

Suppose the invariants hold at layer i. Every coefficient vector in $\mathcal{V}_{i+1}$ has the form $uM_j^{(i)}$ for some $u \in \mathcal{V}_i$ and $j \in [n]$. By the polynomial spanning invariant for $\mathcal{V}_i$, it is therefore a left $\mathbb{F}\langle X_2 \rangle$-linear combination of the rows used to form C_{i+1}. The factorization $C_{i+1} = UV$ then expresses every such row as a left polynomial combination of rows of V, proving (i) for $\mathcal{B}_{i+1}$. Over the free skew field, V has $r = \operatorname{ncrank}(C_{i+1})$ independent rows, while the row space of C_{i+1} has dimension r and is contained in the row span of V; hence these spans are equal. Together with the inductive hypothesis, this proves (ii).

Final step. The polynomial $A[s,t]$ is zero if and only if every coefficient of every monomial is zero. This holds if and only if all coefficient vectors in $\mathcal{V}_d$ have a zero entry at coordinate t. By the invariant, this is equivalent to all vectors in $\mathcal{B}_d$ having a zero entry at coordinate t.

Complexity. At each layer, $|\mathcal{B}_i| \leq w_i \leq s$. Each iteration performs a polynomially many matrix-vector multiplications and one noncommutative rank computation. Over $\mathbb{F}$, deterministic noncommutative-rank algorithms and the factorization procedure below give overall running time polynomial in s. It is important to note that identity testing for partially commutative ABP with two buckets is equivalent to testing the equivalence of two tape deterministic automata. A polynomial-time algorithm was already given for this problem Arvind, Chatterjee and Mukhopadhyay [ACM24] over $\mathbb{Q}$ when the number of buckets is constant.

3.3 Proof of the Factorization Lemma

In this section, we prove Lemma 3.5. For that, we need to use some algebraic tools that we will develop in the following subsections. Throughout this section, unless stated otherwise, $X = \{x_1, \ldots, x_n\}$ denotes a set of noncommutative variables.

3.3.1 Linear Matrix Factorization

For a matrix of noncommutative polynomials, invertibility is defined over the universal skew field of fractions of the noncommutative polynomial ring. The noncommutative rank is the size of a largest invertible sub-matrix. We use the following characterization of the noncommutative rank due to Cohn [Coh95].

Theorem 3.6. (Noncommutative rank). *Let $M \in \mathbb{F}\langle X \rangle^{\ell \times m}$ be a matrix of noncommutative polynomials of noncommutative rank $r < \min\{\ell, m\}$. Then there exist matrices $P \in \mathbb{F}\langle X \rangle^{\ell \times r}$ and $Q \in \mathbb{F}\langle X \rangle^{r \times m}$ such that $M = P \cdot Q$.*

In our proof, we need the following lemmas.

Lemma 3.7. (Linear Matrix Factorization). *Let $\mathbb{F}$ be a field. and $X = \{x_1, x_2, \ldots, x_n\}$. Let $L \in \mathbb{F}\langle X \rangle^{\ell \times m}$ be a noncommutative linear matrix of rank $r < \min\{\ell, m\}$, where "linear" allows a*

constant term (equivalently, the entries are affine linear forms). Then we can construct linear matrices $L_1 \in \mathbb{F}\langle X\rangle^{\ell\times r}$ *and* $L_2 \in \mathbb{F}\langle X\rangle^{r\times m}$ *in deterministic* $\mathrm{poly}(\ell, m, n)$ *time such that* $L = L_1 \cdot L_2$.

The proof of this lemma is implicit from the work of [Coh95, FR04, IQS17, IQS18]. For the sake of completeness, we add a detailed proof in Appendix C.

Lemma 3.8 (Matrix Product Trivialization [AJ22b], [Coh00], PP. 198)**.** *Let* $A \in \mathbb{F}\langle X\rangle^{m\times n}, B \in \mathbb{F}\langle X\rangle^{n\times k}$ *be two matrix polynomial with* A *being linear matrix, such that* $A \cdot B = 0$*. Then there exists an unit* [11] $P \in \mathbb{F}\langle X\rangle^{n\times n}$ *such that for every* $i \in [n]$*, either the* i*-th column of* AP *is all* 0 *or the* i*-th row of* $P^{-1}B$ *is all* 0*.*

Lemma 3.9 (Lemma 2.4, [AJ22a])**.** *Let* $C \in \mathbb{F}\langle X\rangle^{k\times \ell}$ *be a linear matrix and* $D \in \mathbb{F}\langle X\rangle^{\ell\times m}$ *be noncommutative polynomial matrix whose entries are computable by size* s *ABPs, such that* $CD = 0$*. Then in deterministic* $\mathrm{poly}(s, k, \ell, m)$ *time we can compute a unit* $N \in \mathbb{F}\langle X\rangle^{\ell\times\ell}$ *such that,*

1. *For every* $i \in [\ell]$*, either the* i*-th column of* CN *is all* 0*, or the* i*-th row of* $N^{-1}D$ *is all* 0*.*
2. *Each entry of* N *is a polynomial of degree at most* ℓ^2 *and computable by a* poly *size ABP, and also each entry of* N^{-1} *is computable by a* poly *size ABP.*

3.3.2 Higman's Linearization for a Matrix of Noncommutative ABPs

In this section, we bound the ABP complexity of Higman linearization for a rectangular matrix whose entries are computed by noncommutative ABPs. It then suffices to prove the factorization lemma for linear matrices only.

ABPs as Products of Linear Matrices. A standard way to view a noncommutative ABP of size S is as an iterated product of linear matrices. Concretely, if a polynomial $f \in \mathbb{F}\langle X\rangle$ is computed by an ABP of size S, then there exist integers $d \le S$, widths $w_1, \ldots, w_{d+1} \le S$, linear matrices $A_\ell \in \mathbb{F}\langle X\rangle^{w_\ell \times w_{\ell+1}}$, and vectors $u \in \mathbb{F}^{w_1}, v \in \mathbb{F}^{w_{d+1}}$ such that

$$f \;=\; u^\top A_1 A_2 \cdots A_d\, v. \tag{3.10}$$

(Here "linear" means each entry is an affine linear form in X.)

Absorbing $u^\top$ into the first factor and v into the last factor, we may take $w_1 = w_{d+1} = 1$. Thus, for a matrix $M \in \mathbb{F}\langle X\rangle^{s\times s'}$ whose scalar entries have ABP size at most S, after possibly padding the internal widths (by inserting identity blocks), we may assume that

$$M[i,j] \;=\; A_1^{(i,j)} A_2^{(i,j)} \cdots A_d^{(i,j)}, \tag{3.11}$$

[11] A noncommutative polynomial Matrix $C \in \mathbb{F}\langle X\rangle^{\ell\times\ell}$ is called unit, if C is invertible and the inverse of C is again a polynomial matrix.

where $A_1^{(i,j)}$ has one row, $A_d^{(i,j)}$ has one column, and for each $\ell \in [d]$ the factor $A_\ell^{(i,j)}$ is a linear matrix of dimension $w_\ell \times w_{\ell+1}$ with $w_1 = w_{d+1} = 1$ and $w_\ell \le S$, and the width profile $(w_1, \dots, w_{d+1})$ is *independent of* (i, j). This is the model we linearize.

Such a result was previously known only when each entry of the matrix is computed by a *noncommutative formula* [GGdOW20, Proposition A.2].

Higman Linearization for Each Entry

We first isolate the basic gadget that linearizes a single product entry while introducing only upper/lower triangular multipliers with diagonal entries that are 1.

Lemma 3.12 (Higman linearization for one product block [AJ25]). *Let $B \in \mathbb{F}\langle X\rangle^{w_1 \times w_{d+1}}$ be a block polynomial of the form*

$$B = A_1 A_2 \cdots A_d,$$

where $A_\ell \in \mathbb{F}\langle X\rangle^{w_\ell \times w_{\ell+1}}$ is a linear matrix. *Let $w := \sum_{\ell=2}^{d} w_\ell$. Then there exist matrices P_B, Q_B, L_B such that*

$$\begin{pmatrix} B & 0 \\ 0 & I_w \end{pmatrix} = P_B \cdot L_B \cdot Q_B, \tag{3.13}$$

where:

1. *P_B is upper triangular and Q_B is lower triangular,*
2. *all diagonal entries of P_B, Q_B are equal to 1,*
3. *L_B is a* linear matrix *(every entry is either 0, 1, or an affine linear form in X),*
4. *every nonzero entry of P_B, Q_B is either ± 1 or an entry of one of the factors A_ℓ. Hence, if each A_ℓ is computable by* poly(S)*-size ABPs, then so are the entries of P_B, Q_B.*

Moreover, L_B can be taken in the companion form

$$L_B = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & A_1 \\ -A_d & I_{w_d} & 0 & \cdots & 0 & 0 \\ 0 & -A_{d-1} & I_{w_{d-1}} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & I_{w_3} & 0 \\ 0 & 0 & 0 & \cdots & -A_2 & I_{w_2} \end{pmatrix}. \tag{3.14}$$

Proof. This is the standard Higman linearization (sometimes called "Higman's trick"). One can prove (3.13) by repeatedly applying the $d = 2$ elimination step: for any compatible linear matrices

$U \in \mathbb{F}\langle X\rangle^{a\times b}$ and $V \in \mathbb{F}\langle X\rangle^{b\times c}$, set

$$\begin{pmatrix} UV & 0 \\ 0 & I_b \end{pmatrix} = \underbrace{\begin{pmatrix} I_a & -U \\ 0 & I_b \end{pmatrix}}_{\text{upper triangular}} \cdot \underbrace{\begin{pmatrix} 0 & U \\ -V & I_b \end{pmatrix}}_{\text{linear}} \cdot \underbrace{\begin{pmatrix} I_c & 0 \\ V & I_b \end{pmatrix}}_{\text{lower triangular}} .$$

after a suitable block permutation to match dimensions (equivalently, embed into the bottom-right corner). Iterating this construction for the product $A_1 \cdots A_d$ yields the companion matrix (3.14), and the accumulated left and right multipliers remain upper/lower triangular with diagonal 1. Every nonzero off-diagonal entry introduced is a copy of some A_ℓ or a constant; hence, it is ABP-computable whenever the A_ℓ are. □

Higman Linearization for a Rectangular Matrix of ABP Entries

We now lift the local gadget to a full $s \times s'$ matrix M, linearizing each entry in place while preserving the global triangular structure.

Lemma 3.15 (Higman linearization for a matrix of ABP-entries)**.** *Let $M \in \mathbb{F}\langle X\rangle^{s\times s'}$ be a rectangular matrix such that for every $(i,j) \in [s] \times [s']$*

$$M[i,j] \;=\; A_1^{(i,j)} A_2^{(i,j)} \cdots A_d^{(i,j)}, \tag{3.16}$$

where for each $\ell \in [d]$, the factor $A_\ell^{(i,j)}$ is a linear matrix of dimension $w_\ell \times w_{\ell+1}$ (the width profile $(w_1, \ldots, w_{d+1})$ is independent of (i,j)). Let

$$w := \sum_{\ell=2}^{d} w_\ell, \qquad t := w \cdot ss'.$$

Then there exist matrices P, L, Q over $\mathbb{F}\langle X\rangle$ such that

$$M \oplus I_t \;=\; P \cdot L \cdot Q, \tag{3.17}$$

with the following properties:

1. *P is (block) upper triangular and Q is (block) lower triangular,*
2. *all diagonal entries of P, Q are equal to 1,*
3. *L is a* linear matrix *(all its entries are affine linear forms in X),*
4. *every nonzero entry of P, Q is computable by an ABP of size* $\mathrm{poly}(s, s', d, w_1, \ldots, w_{d+1})$.

Proof. We linearize the entries of M one-by-one using Lemma 3.12 as a block gadget, and we do so in an order that preserves the triangularity of the accumulated multipliers.

Order of Linearization. Fix the lexicographic order scanning columns from right to left, and within each column from bottom to top:

$$(s, s'), (s-1, s'), \ldots, (1, s'), (s, s'-1), \ldots, (1, 1).$$

Invariant. After processing τ entries, we maintain a matrix $\widetilde{M}$ of size

$$(s \cdot w_1 + w \cdot \tau) \times (s' \cdot w_{d+1} + w \cdot \tau)$$

such that:

- all processed positions have been replaced by linear companion blocks,
- all unprocessed positions still contain their original product blocks,
- $\widetilde{M}$ is obtained from M by adjoining τ slack identity blocks I_w and multiplying on the left and right by block upper/lower triangular matrices with diagonal 1.

One Step: Local Linearization. Let (i, j) be the next unprocessed position. We adjoin one additional slack block I_w, increasing τ to $\tau + 1$.

We perform the Higman linearization locally on the coordinates corresponding to:

- the block $M[i, j]$, and
- the newly added slack block I_w.

Within these coordinates, the matrix contains a sub-block of the form

$$\begin{pmatrix} M[i,j] & 0 \\ 0 & I_w \end{pmatrix}.$$

Applying Lemma 3.12 to this sub-block yields matrices $P_{i,j}, Q_{i,j}$ and a linear block $L_{i,j}$ satisfying

$$\begin{pmatrix} M[i,j] & 0 \\ 0 & I_w \end{pmatrix} = P_{i,j} \cdot L_{i,j} \cdot Q_{i,j}.$$

We extend $P_{i,j}$ and $Q_{i,j}$ to the full matrix by acting as identity outside these coordinates. Denote the resulting matrices by $\widehat{P}_{i,j}$ and $\widehat{Q}_{i,j}$.

Since $P_{i,j}$ is upper triangular and $Q_{i,j}$ is lower triangular with diagonal entries equal to 1, their extensions $\widehat{P}_{i,j}$ and $\widehat{Q}_{i,j}$ are block upper and lower triangular with diagonal 1. Thus we obtain

$$\widetilde{M} \oplus I_w = \widehat{P}_{i,j} \cdot \widetilde{M}^{\,\mathrm{new}} \cdot \widehat{Q}_{i,j}, \tag{3.18}$$

where $\widetilde{M}^{\,\mathrm{new}}$ agrees with $\widetilde{M}$ away from the selected block and the fresh slack coordinates. This follows from the triangular form of the local multipliers: the newly added slack row and column

are zero outside the fresh identity block, so the row and column operations do not alter the other original entries.

Completion. Repeating this process for all ss' entries, we obtain a matrix L in which every entry has been linearized. Let P and Q be the products of all accumulated multipliers.

Since each step multiplies by block upper (resp. lower) triangular matrices with diagonal 1, and this property is preserved under multiplication, the final matrices P, Q satisfy the required triangularity conditions. The total number of slack blocks used is $t = w \cdot ss'$.

ABP Size Bound. Each step introduces only constants and entries from the matrices $A_\ell^{(i,j)}$ into P and Q. Hence every entry of P, Q is computable by an ABP of size polynomial in $(s, s', d, w_1, \ldots, w_{d+1})$. □

Remark 3.19. *The above construction avoids any row or column permutations. All steps are implemented via local Higman linearizations embedded into the matrix using slack identity blocks. Consequently, all left multipliers are upper triangular with diagonal* 1*, and all right multipliers are lower triangular with diagonal* 1*.* ◊

If the entries of M are given directly by ABPs of size $\leq S$, then by the standard conversion (3.10) we can rewrite each entry as a product of $d \leq S$ linear matrices with widths at most S. After padding to equalize width profiles across entries, Lemma 3.15 applies and yields $t = \mathrm{poly}(s, s', S)$.

Corollary 3.20 (Higman linearization for ABP entries)**.** *Let $M \in \mathbb{F}\langle X\rangle^{s\times s'}$ be such that each entry is computed by a noncommutative ABP of size at most S. Then there exists $t = \mathrm{poly}(s, s', S)$ and matrices P, L, Q such that*

$$M \oplus I_t = P \cdot L \cdot Q,$$

where P is upper triangular with diagonal 1*, Q is lower triangular with diagonal* 1*, L is linear, and every entry of P, Q is computable by* $\mathrm{poly}(s, s', S)$*-size ABPs.*

3.3.3 Proof of Factorization Lemma

Finally, we prove the existence of *small* factors witnessing the *low rank*.

Lemma 3.5. (Factorization Lemma) *Let $M \in \mathbb{F}\langle X\rangle^{\ell\times m}$ be a matrix of noncommutative polynomials with* $\mathrm{ncrank}(M) = r < \min\{\ell, m\}$*. Assume each entry of M is computed by a size s noncommutative ABP. Then we can compute a nontrivial factorization of M, namely $M = G \cdot H$, such that the entries of $G \in \mathbb{F}\langle X\rangle^{\ell\times r}$ and $H \in \mathbb{F}\langle X\rangle^{r\times m}$ have* $\mathrm{poly}(s, \ell, m)$*-size noncommutative ABPs.*

Proof. Let $M \in \mathbb{F}\langle X\rangle^{\ell\times m}$ with $\mathrm{ncrank}(M) = r < \min\{m, \ell\}$. We first apply Corollary 3.20 to M and obtain

$$\left[\begin{array}{c|c} M & 0 \\ \hline 0 & I_{k'} \end{array}\right] = P \cdot L \cdot D, \tag{3.21}$$

where $L \in \mathbb{F}\langle X\rangle^{(\ell+k')\times(m+k')}$ is a linear matrix, and $P \in \mathbb{F}\langle X\rangle^{(\ell+k')\times(\ell+k')}$, $D \in \mathbb{F}\langle X\rangle^{(m+k')\times(m+k')}$ are upper triangular and lower triangular full rank matrices, respectively, with all diagonal entries equal to 1. Moreover, every entry of P, L, D is computable by a size $\mathrm{poly}(s)$ noncommutative ABP.

Our goal is to compute a nontrivial rank factorization $M = G \cdot H$ with inner rank r, such that the entries of G and H have $\mathrm{poly}(s)$-size ABPs.

Write the block decompositions $P = \left[\begin{array}{c|c} P_{1,1} & P_{1,2} \\ \hline 0 & P_{2,2} \end{array}\right]$, $D = \left[\begin{array}{c|c} D_{1,1} & 0 \\ \hline D_{2,1} & D_{2,2} \end{array}\right]$. Since P and D are invertible units (they are triangular matrices with 1's on the diagonal), multiplication by P or D preserves noncommutative rank. Moreover, noncommutative rank is additive under direct sums. Hence, taking noncommutative rank on both sides of Equation 3.21 gives $\mathrm{ncrank}(L) = \mathrm{ncrank}\left(\left[\begin{array}{c|c} M & 0 \\ \hline 0 & I_{k'} \end{array}\right]\right) = \mathrm{ncrank}(M) + k' = r + k' = r'$(say).

Using Lemma 3.7, we first factorize L,

$$\underbrace{L}_{(\ell+k')\times(m+k')} = \underbrace{L_1}_{(\ell+k')\times r'} \cdot \underbrace{L_2}_{r'\times(m+k')}. \tag{3.22}$$

Since the inner dimension of the factorization is exactly $\mathrm{ncrank}(L) = r'$, both L_1 and L_2 have full noncommutative column rank and row rank, respectively. In particular, $\mathrm{ncrank}(L_1) = \mathrm{ncrank}(L_2) = r'$.

Moreover, we assume the following block decomposition of L_1 and L_2.

$$\begin{aligned} L_1 &= \left[\begin{array}{c} L_{1,1} \\ \hline L_{1,2} \end{array}\right] \quad L_{1,1} \in \mathbb{F}\langle X\rangle^{\ell\times r'},\ L_{1,2} \in \mathbb{F}\langle X\rangle^{k'\times r'} \\ L_2 &= \left[\begin{array}{c|c} L'_{1,1} & L'_{1,2} \end{array}\right] \quad L'_{1,1} \in \mathbb{F}\langle X\rangle^{r'\times m},\ L'_{1,2} \in \mathbb{F}\langle X\rangle^{r'\times k'}. \end{aligned} \tag{3.23}$$

Combining Equation 3.22 and Equation 3.23, we have the following equality.

$$\begin{aligned} \left[\begin{array}{c|c} M & 0 \\ \hline 0 & I_{k'} \end{array}\right] &= \left[\begin{array}{c|c} P_{1,1} & P_{1,2} \\ \hline 0 & P_{2,2} \end{array}\right] \left[\begin{array}{c} L_{1,1} \\ \hline L_{1,2} \end{array}\right] \times \left[\begin{array}{c|c} L'_{1,1} & L'_{1,2} \end{array}\right] \left[\begin{array}{c|c} D_{1,1} & 0 \\ \hline D_{2,1} & D_{2,2} \end{array}\right] \\ &= \left[\begin{array}{c} P_{1,1}L_{1,1} + P_{1,2}L_{1,2} \\ \hline P_{2,2}L_{1,2} \end{array}\right] \times \left[\begin{array}{c|c} L'_{1,1}D_{1,1} + L'_{1,2}D_{2,1} & L'_{1,2}D_{2,2} \end{array}\right]. \end{aligned} \tag{3.24}$$

Consider the $(1,1)$-th block of both matrices in Equation 3.24.

$$\underbrace{M}_{\ell\times m} = \underbrace{(P_{1,1}L_{1,1} + P_{1,2}L_{1,2})}_{\ell\times r'} \times \underbrace{(L'_{1,1}D_{1,1} + L'_{1,2}D_{2,1})}_{r'\times m}. \tag{3.25}$$

Similarly, equating the $(2,1)$-th block and $(2,2)$-th block from Equation 3.24 we have

$$0 = P_{2,2}(L_{1,2})(L'_{1,1}D_{1,1} + L'_{1,2}D_{2,1}) \tag{3.26}$$

and

$$I_{k'} = P_{2,2}(L_{1,2})(L'_{1,2}D_{2,2}). \tag{3.27}$$

Consider Equation 3.27. Since $P_{2,2}$ and $D_{2,2}$ are invertible units, we have $k' \leq \text{ncrank}(L_{1,2})$. But $L_{1,2}$ has only k' rows. Hence, $L_{1,2}$ has full noncommutative row rank, i.e., $\text{ncrank}(L_{1,2}) = k'$.

Consider now Equation 3.26. Since $P_{2,2}$ is invertible, $L_{1,2}(L'_{1,1}D_{1,1} + L'_{1,2}D_{2,1}) = 0$. Let $E := L'_{1,1}D_{1,1} + L'_{1,2}D_{2,1} \in \mathbb{F}\langle X\rangle^{r'\times m}$. Since $\text{ncrank}(L_{1,2}) = k'$, the matrix $L_{1,2}$ has full row rank over the free skew field. Therefore its right kernel has dimension $r' - k' = r$. Moreover, from $L_{1,2}E = 0$, every column of E belongs to the right kernel of $L_{1,2}$. Consequently, $\text{ncrank}(E) \leq r$. On the other hand, from Equation 3.25, we have $\text{ncrank}(M) = r \leq \text{ncrank}(E)$. Hence, $\text{ncrank}(E) = r$.

We now have $\underbrace{(L_{1,2})}_{k'\times r'}\underbrace{(L'_{1,1}D_{1,1} + L'_{1,2}D_{2,1})}_{r'\times m} = 0$ with $\text{ncrank}(L_{1,2}) = k'$ and $\text{ncrank}(L'_{1,1}D_{1,1} + L'_{1,2}D_{2,1}) = r$.

Since $L_{1,2}$ is a linear matrix, we can apply the matrix product trivialization, Lemma 3.9. Hence, there exists a unit matrix N whose entries, as well as the entries of N^{-1}, are computable by polynomial-size noncommutative ABPs, such that for every $a \in [r']$, either the a-th column of $L_{1,2}N$ is zero or the a-th row of $N^{-1}(L'_{1,1}D_{1,1} + L'_{1,2}D_{2,1})$ is zero.

We next show that the product trivialization separates exactly k' coordinates corresponding to $L_{1,2}$ and r coordinates corresponding to $L'_{1,1}D_{1,1} + L'_{1,2}D_{2,1}$. Let $S \subseteq [r']$ denote the indices of the nonzero columns of $L_{1,2}N$, and let $T \subseteq [r']$ denote the indices of the nonzero rows of $N^{-1}(L'_{1,1}D_{1,1} + L'_{1,2}D_{2,1})$. By the product trivialization property, $S \cap T = \varnothing$.

Since N is a unit, $\text{ncrank}(L_{1,2}N) = \text{ncrank}(L_{1,2}) = k'$. Thus, $L_{1,2}N$ must have at least k' nonzero columns, and hence $|S| \geq k'$. Similarly, $\text{ncrank}\Big(N^{-1}(L'_{1,1}D_{1,1} + L'_{1,2}D_{2,1})\Big) = r$, so this matrix must have at least r nonzero rows, and hence $|T| \geq r$.

Since S and T are disjoint subsets of $[r']$ and $r' = r + k'$, we have $r' = r + k' \leq |S| + |T| \leq r'$. Therefore, $|S| = k'$, $|T| = r$, and $S \sqcup T = [r']$.

Hence, there exists a permutation matrix Π such that the last k' columns of $L_{1,2}N\Pi$ are nonzero while the first r columns are zero, and the first r rows of $\Pi^{-1}N^{-1}(L'_{1,1}D_{1,1} + L'_{1,2}D_{2,1})$ are nonzero while the remaining k' rows are zero. In particular, $\Pi^{-1}N^{-1}(L'_{1,1}D_{1,1} + L'_{1,2}D_{2,1}) = \left[\begin{array}{c} H \\ \hline 0 \end{array}\right]$ for some $H \in \mathbb{F}\langle X\rangle^{r\times m}$. Since the matrix on the left has noncommutative rank r, it follows that $\text{ncrank}(H) = r$.

From Equation 3.25 we have the following factorization,

$$\begin{aligned} M &= ((P_{1,1}L_{1,1} + P_{1,2}L_{1,2})N\Pi)\left(\Pi^{-1}N^{-1}(L'_{1,1}D_{1,1} + L'_{1,2}D_{2,1})\right) \\ &= \left[\begin{array}{c|c} \underbrace{G}_{\ell\times r} & \underbrace{G'}_{\ell\times k'} \end{array}\right]\left[\begin{array}{c} H \\ \hline 0 \end{array}\right] = \underbrace{G}_{\ell\times r}\cdot\underbrace{H}_{r\times m}. \end{aligned} \tag{3.28}$$

Since $\operatorname{ncrank}(M) = r$ and the inner dimension of this factorization is r, both factors have full rank in the corresponding direction. In particular, $\operatorname{ncrank}(G) = \operatorname{ncrank}(H) = r$.

ABP-size bound. It remains to show that every entry of G and H is computable by a polynomial-size noncommutative ABP.

By Corollary 3.20, every entry of P and D is computable by a $\operatorname{poly}(s)$-size noncommutative ABP. The matrices L_1 and L_2 obtained from Lemma 3.7 are linear matrices. Hence, all entries of $L_{1,1}, L_{1,2}, L'_{1,1}$ and $L'_{1,2}$ have constant-size ABPs.

It follows that every entry of $P_{1,1}L_{1,1} + P_{1,2}L_{1,2}$ is computable by a polynomial-size noncommutative ABP. Indeed, each entry is a sum of at most polynomially many products of entries of P with linear forms.

Similarly, every entry of $L'_{1,1}D_{1,1} + L'_{1,2}D_{2,1}$ is computable by a polynomial-size noncommutative ABP. Here again, each entry is a sum of polynomially many products of a linear form with an entry of D.

By Lemma 3.9, the entries of the unit matrix N and its inverse N^{-1} are computable by polynomial-size noncommutative ABPs. Multiplication by the permutation matrix Π or Π^{-1} only permutes rows or columns and therefore introduces no additional algebraic complexity.

Consequently, every entry of $(P_{1,1}L_{1,1} + P_{1,2}L_{1,2})N\Pi$ and every entry of $\Pi^{-1}N^{-1}(L'_{1,1}D_{1,1} + L'_{1,2}D_{2,1})$ is computable by a polynomial-size noncommutative ABP. The dimensions of all matrices involved are polynomially bounded in the input size by the Higman linearization construction, so the polynomially many additions and multiplications required for these matrix products preserve polynomial ABP size.

Finally, G and H are sub-matrices of the above two matrices. Hence every entry of G and H is computable by a $\operatorname{poly}(s)$-size noncommutative ABP. This completes the proof. □

We have the following corollary, which is required to prove Lemma 1.6.

Corollary 3.29. *Let $A_1, \ldots, A_n \in \mathbb{F}\langle X\rangle^{k\times m}$, with every entry computed by a size-s noncommutative ABP. Let*

$$A = \begin{pmatrix} A_1 \\ \vdots \\ A_n \end{pmatrix} \in \mathbb{F}\langle X\rangle^{(nk)\times m}, \qquad r = \operatorname{ncrank}(A).$$

Then there exist a matrix $B \in \mathbb{F}\langle X\rangle^{r\times m}$ and matrices $C_i \in \mathbb{F}\langle X\rangle^{k\times r}$ such that $A_i = C_iB$ for every

$i \in [n]$. The matrix B has full noncommutative row rank, the stacked matrix $C = (C_1^\top, \ldots, C_n^\top)^\top$ has full noncommutative column rank, and all entries of B and the C_i are computable by $\mathrm{poly}(s, n, k, m)$*-size noncommutative ABPs.*

Proof. If $r < \min\{nk, m\}$, apply Lemma 3.5 to obtain a rank factorization $A = CB$ with inner dimension r, and partition C into n consecutive blocks $C_1, \ldots, C_n$ of k rows each. If $r = nk \leq m$, take $C = I_{nk}$ and $B = A$. If $r = m \leq nk$, take $C = A$ and $B = I_m$. These choices also cover the square full-rank case.

In every case, $A = CB$ and the inner dimension is $r = \mathrm{ncrank}(A)$. Hence B has full noncommutative row rank and C has full noncommutative column rank: otherwise the rank of the product would be smaller than r. In particular, C admits a left inverse over the free skew field. The stated ABP-size bounds follow from Lemma 3.5 in the non-full-rank case and are immediate in the boundary cases. □

3.4 Construction of witness matrices

We end with a proof of Lemma 1.6 using the machinery we have developed so far.

Lemma 1.6. (Witness Theorem for Partially Commutative ABP Identities) *Let F be a field and A be a partially commutative homogeneous (over X_1) ABP in $\mathbb{F}\langle X_1 \sqcup X_2\rangle$ of size s computing the identically zero polynomial. Let d be the depth of A, and let w_i denote the width of the i-th layer. Then there exist $2d-1$ matrices $\{\lambda'_0, \ldots, \lambda'_{d-1}\}$ and $\{T'_1, \ldots, T'_{d-1}\}$ such that*

$$\lambda'_i \in \mathbb{F}^{w_i \times w_i}\langle X_2\rangle, \qquad T'_i \in \mathbb{F}^{w_{i-1} \times w_i}\langle X_1 \sqcup X_2\rangle,$$

and the following hold:

$$\lambda'_{d-i} \cdot \overline{A}_i = \overline{0} \quad \forall\, i = 1, 2, \ldots, d, \tag{1.6}$$

$$\lambda'_{d-i} \cdot \overline{A}_i = T'_{d-i+1} \cdot \lambda'_{d-i+1} \cdot \overline{A}_{i-1} \quad \forall\, i = 2, 3, \ldots, d, \tag{1.7}$$

$$\lambda'_0 := 1, \qquad \lambda'_0 \cdot \overline{A}_d = \overline{A}_d = 0. \tag{1.8}$$

Moreover, every entry of the matrices λ'_i and T'_j is computable by a polynomial-size noncommutative ABP.

Proof. Let A be a partially commutative ABP over $X = X_1 \sqcup X_2$ of depth d and layer widths $w_0, \ldots, w_d$, which is *homogeneous over* X_1 and computes the identically zero polynomial in $\mathbb{F}\langle X_1 \sqcup X_2\rangle$. For $i \in [d]$, recall the column vector

$$\overline{A}_i \;=\; \big(A[u_{d-i,1}, t],\; A[u_{d-i,2}, t],\; \ldots,\; A[u_{d-i,w_{d-i}}, t]\big)^\top \in \mathbb{F}\langle X_1 \sqcup X_2\rangle^{w_{d-i}}.$$

In particular, $\overline{A}_d$ is the polynomial computed by A, and hence $\overline{A}_d = 0$.

Homogeneity Convention. Since A is homogeneous over X_1 from our assumption, we view each edge label of A as a homogeneous linear form in the X_1-variables whose coefficients lie in $\mathbb{F}\langle X_2\rangle$. Equivalently, for every layer transition $i \to i+1$ there exist coefficient matrices

$$M_1, \ldots, M_n \in \mathbb{F}\langle X_2\rangle^{w_i \times w_{i+1}}$$

such that the one-layer ABP recurrence from layer i to layer $i+1$ can be written as

$$\overline{A}_{d-i} \;=\; \sum_{j=1}^{n} x_{1,j} \cdot \big(M_j \cdot \overline{A}_{d-i-1}\big), \tag{3.30}$$

where the product is matrix-vector multiplication over $\mathbb{F}\langle X_2\rangle$, and the $x_{1,j}$ appears as *left* multipliers.

Step 0: initialize λ'_0. Set

$$\lambda'_0 \;:=\; 1 \in \mathbb{F}\langle X_2\rangle^{w_0 \times w_0}.$$

Then $\lambda'_0 \cdot \overline{A}_d = \overline{A}_d = 0$, which is Item (3) in the statement.

Inductive construction. We construct, for $i = 0, 1, \ldots, d-2$, matrices

$$T'_{i+1} \in \mathbb{F}\langle X_1 \sqcup X_2\rangle^{w_i \times w_{i+1}} \qquad \text{and} \qquad \lambda'_{i+1} \in \mathbb{F}\langle X_2\rangle^{w_{i+1} \times w_{i+1}}$$

such that for every i we have both:

$$\lambda'_i \cdot \overline{A}_{d-i} = 0, \tag{3.31}$$

$$\lambda'_i \cdot \overline{A}_{d-i} = T'_{i+1} \cdot \lambda'_{i+1} \cdot \overline{A}_{d-i-1}. \tag{3.32}$$

Note that Equation 3.32 is exactly the desired transition identity, and together with Equation 3.31 it will imply the next vanishing statement $\lambda'_{i+1} \cdot \overline{A}_{d-i-1} = 0$.

Fix $i \in \{0, \ldots, d-2\}$ and assume the inductive hypothesis

$$\lambda'_i \cdot \overline{A}_{d-i} = 0. \tag{3.33}$$

Step 1: form the rectangular coefficient matrix over $\mathbb{F}\langle X_2\rangle$. Using Equation 3.30 and left-multiplying by λ'_i, we get

$$\lambda'_i \cdot \overline{A}_{d-i} \;=\; \sum_{j=1}^{n} x_{1,j} \cdot \big((\lambda'_i M_j) \cdot \overline{A}_{d-i-1}\big). \tag{3.34}$$

By Equation 3.33, the left-hand side is 0. Since the variables in X_1 are noncommuting and appear as *leftmost* multipliers in Equation 3.34, homogeneity over X_1 implies that each coefficient must vanish:

$$(\lambda_i' M_j) \cdot \overline{A}_{d-i-1} = 0 \qquad \forall\, j \in [n]. \tag{3.35}$$

Define the stacked rectangular matrix

$$T \;:=\; \begin{pmatrix} \lambda_i' \cdot M_1 \\ \lambda_i' \cdot M_2 \\ \vdots \\ \lambda_i' \cdot M_n \end{pmatrix} \;\in\; \mathbb{F}\langle X_2 \rangle^{(nw_i) \times w_{i+1}}. \tag{3.36}$$

Then Equation 3.35 is equivalent to

$$T \cdot \overline{A}_{d-i-1} = 0. \tag{3.37}$$

Step 2: apply Corollary 3.29 to extract a common left factor. Apply Corollary 3.29 over the ring $\mathbb{F}\langle X_2 \rangle$ to the matrices $\lambda_i' M_1, \ldots, \lambda_i' M_n \in \mathbb{F}\langle X_2 \rangle^{w_i \times w_{i+1}}$. We obtain:

- a basis matrix $B \in \mathbb{F}\langle X_2 \rangle^{w'_{i+1} \times w_{i+1}}$,
- and coefficient matrices $C_1, \ldots, C_n \in \mathbb{F}\langle X_2 \rangle^{w_i \times w'_{i+1}}$,

such that for every $j \in [n]$,

$$\lambda_i' \cdot M_j \;=\; C_j \cdot B \qquad \forall\, j \in [n]. \tag{3.38}$$

Step 3: define λ'_{i+1} and T'_{i+1}. Note that it may happen that $w'_{i+1} < w_{i+1}$. In that case, add 0 rows to make the number of rows of B equal to w_{i+1}. Let this new matrix be B'. Similarly, add 0 columns to each C_j to obtain the number of columns equal to w_{i+1}. Let these new matrices be C'_j.

$$\lambda'_{i+1} \;:=\; B' \in \mathbb{F}\langle X_2 \rangle^{w_{i+1} \times w_{i+1}}, \qquad T'_{i+1} \;:=\; \sum_{j=1}^{n} x_{1,j} \cdot C'_j \in \mathbb{F}\langle X_1 \sqcup X_2 \rangle^{w_i \times w_{i+1}}.$$

Then, combining Equation 3.30 with Equation 3.38, we obtain the propagation identity:

$$\begin{aligned} \lambda_i' \cdot \overline{A}_{d-i} &= \sum_{j=1}^{n} x_{1,j} \cdot \left((\lambda_i' M_j) \cdot \overline{A}_{d-i-1} \right) = \sum_{j=1}^{n} x_{1,j} \cdot \left((C_j B) \cdot \overline{A}_{d-i-1} \right) \\ &= \left(\sum_{j=1}^{n} x_{1,j} \cdot C_j \right) \cdot \left(B \cdot \overline{A}_{d-i-1} \right) = T'_{i+1} \cdot \lambda'_{i+1} \cdot \overline{A}_{d-i-1}. \end{aligned}$$

Padding by zero rows and columns does not change the product $C_j B$, so the last equality is valid with C'_j and B'. This proves Equation 3.32 for the current i.

Step 4: derive the next vanishing statement $\lambda'_{i+1} \cdot \overline{A}_{d-i-1} = 0$. We now derive the next vanishing statement, namely $\lambda'_{i+1} \cdot \overline{A}_{d-i-1} = 0$. We have two cases to consider here. The first case is when the stacked matrix T, defined in Equation 3.36 has noncommutative rank $r < w_{i+1}$. For this case, recall from Equation 3.35; for every $j \in [n]$,

$$(\lambda'_i M_j) \cdot \overline{A}_{d-i-1} = 0.$$

In Step 2, we applied Corollary 3.29 to the family $\{\lambda'_i M_j\}_{j\in[n]}$ (each of dimension $w_i \times w_{i+1}$) and obtained a factorization of the form

$$\lambda'_i \cdot M_j \;=\; C_j \cdot B \qquad \forall\, j \in [n], \quad \text{where } B \in \mathbb{F}\langle X_2\rangle^{r\times w_{i+1}}, \text{ and } C_j \in \mathbb{F}\langle X_2\rangle^{w_i \times r}. \tag{3.39}$$

Substituting Equation 3.39 into the coefficient vanishing gives, for all $j \in [n]$,

$$0 \;=\; (\lambda'_i M_j) \cdot \overline{A}_{d-i-1} \;=\; (C_j B) \cdot \overline{A}_{d-i-1} \;=\; C_j \cdot \left(B \cdot \overline{A}_{d-i-1}\right) \forall\, j \in [n].$$

Hence

$$C \cdot \left(B \cdot \overline{A}_{d-i-1}\right) = 0 \quad \text{where} \quad C \;:=\; \begin{pmatrix} C_1 \\ C_2 \\ \vdots \\ C_n \end{pmatrix} \in \mathbb{F}\langle X_2\rangle^{(nw_i)\times r}. \tag{3.40}$$

The Corollary 3.29 shows that the stacked coefficient matrix C has full noncommutative column rank, r. Equivalently, C admits a left inverse over $\mathbb{F}\langle\langle X_2\rangle\rangle$; i.e., there exists a matrix $C^\dagger$ (over the free skew field) such that

$$C^\dagger \cdot C \;=\; I_r.$$

Left-multiplying Equation 3.40 by $C^\dagger$ therefore gives

$$B \cdot \overline{A}_{d-i-1} = 0 \implies B' \cdot \overline{A}_{d-i-1} = 0 \implies \lambda'_{i+1} \cdot \overline{A}_{d-i-1} = 0. \tag{3.41}$$

On the other hand, suppose that the stacked matrix T defined in Equation 3.36 has full noncommutative column rank, i.e., $\mathrm{ncrank}(T) = w_{i+1}$ over $\mathbb{F}\langle\langle X_2\rangle\rangle$. Then, by Equation 3.35, we have $T \cdot \overline{A}_{d-i-1} = 0$. Each entry of $\overline{A}_{d-i-1}$ is a partially commutative ABP that is homogeneous over X_1 with coefficients in $\mathbb{F}\langle X_2\rangle$. Accordingly, we may write $\overline{A}_{d-i-1} = \sum_m m \cdot v_m$, where the sum ranges over homogeneous noncommutative monomials m in X_1 and $v_m \in \mathbb{F}\langle X_2\rangle^{w_{i+1}\times 1}$ is the coefficient vector of m. Substituting this expansion yields

$$0 = T \cdot \overline{A}_{d-i-1} = \sum_m m \cdot (T \cdot v_m). \tag{3.42}$$

Since the variables in X_1 are noncommuting and all monomials m appearing above have the same X_1-degree, the representation $\sum_m m \cdot (T \cdot v_m) = 0$ is unique; hence $T \cdot v_m = 0$ for every m. Since T has full noncommutative column rank, its right kernel over the free skew field is zero. Thus every $v_m = 0$, and therefore $\overline{A}_{d-i-1} = 0$, so trivially $\lambda'_{i+1} \cdot \overline{A}_{d-i-1} = 0$.

Re-indexing to match Equation 1.7–Equation 1.8. Setting $j = d - i$ in Equation 3.39 gives

$$\lambda'_{d-j} \cdot \overline{A}_j = T'_{d-j+1} \cdot \lambda'_{d-j+1} \cdot \overline{A}_{j-1} \qquad \forall\, j = 2, 3, \ldots, d,$$

which is Equation 1.8. Moreover, Equation 3.41 yields Equation 1.7, namely

$$\lambda'_{d-j} \cdot \overline{A}_j = 0 \qquad \forall\, j = 1, 2, \ldots, d.$$

ABP-size of entries. Each entry of $\lambda'_i M_j$ is computed by a size $\mathrm{poly}(s)$ ABPs over $\mathbb{F}\langle X_2 \rangle$, and so are the entries of the stacked matrix T in Equation 3.36. Applying Corollary 3.29 preserves the size $\mathrm{poly}(s)$ of the computability of these entries, and therefore, the padded matrix $\lambda'_{i+1} = B'$ and each C'_j have polynomial size ABPs. Finally,

$$T'_{i+1} = \sum_{j=1}^{n} x_{1,j} \cdot C'_j$$

is a homogeneous linear form in X_1 with ABP-computable coefficients in $\mathbb{F}\langle X_2 \rangle$, again of size $\mathrm{poly}(s)$. This completes the proof of Lemma 1.6. □

Acknowledgement

We acknowledge the use of LLM (GPT-5.6) to improve the language, grammar and formatting the manuscripts.

References

[ACM24] Vikraman Arvind, Abhranil Chatterjee, and Partha Mukhopadhyay. Trading Determinism for Noncommutativity in Edmonds' Problem. In *65th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2024, Chicago, IL, USA, October 27-30, 2024*, pages 539–559. IEEE, 2024.

[AJ22a] V. Arvind and Pushkar S. Joglekar. Matrix Polynomial Factorization via Higman Linearization, 2022. Pre-print available at `arXiv:2203.16978`.

[AJ22b] Vikraman Arvind and Pushkar S. Joglekar. On Efficient Noncommutative Polynomial Factorization via Higman Linearization. In *37th Computational Complexity Conference, CCC 2022, July 20-23, 2022, Philadelphia, PA, USA*, volume 234 of *LIPIcs*, pages 12:1–12:22. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022.

[AJ25] V. Arvind and Pushkar S. Joglekar. On Efficient Noncommutative Polynomial Factorization via Higman Linearization, 2025. Pre-print available at `arXiv:2202.09883`.

[Ajt88] Miklós Ajtai. The Complexity of the Pigeonhole Principle. In *Proceedings of the 29th Annual Symposium on Foundations of Computer Science*, pages 346–355. IEEE Computer Society, 1988.

[Bus87] Samuel R. Buss. Polynomial Size Proofs of the Propositional Pigeonhole Principle. *The Journal of Symbolic Logic*, 52(4):916–927, 1987.

[CF69] P. Cartier and D. Foata. Problémes combinatoires de commutation et réarrangements. *Lecture Notes in Mathematics*, 1969.

[Coh95] P. M. Cohn. *Skew Fields: Theory of General Division Rings*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1995.

[Coh00] Paul M. Cohn. *Introduction to Ring Theory*. Springer Undergraduate Mathematics Series. Springer, 2000.

[Coh06] P. M. Cohn. *Free Ideal Rings and Localization in General Rings*. Cambridge University Press, 2006.

[CR79] Stephen A. Cook and Robert A. Reckhow. The Relative Efficiency of Propositional Proof Systems. *Journal of Symbolic Logic*, 44(1):36–50, 1979.

[FR04] Marc Fortin and Christophe Reutenauer. COMMUTATIVE/NONCOMMUTATIVE RANK OF LINEAR MATRICES AND SUBSPACES OF MATRICES OF LOW RANK. *Séminaire Lotharingien de Combinatoire [electronic only]*, 52, 01 2004.

[FSTW21] Michael A. Forbes, Amir Shpilka, Iddo Tzameret, and Avi Wigderson. Proof Complexity Lower Bounds from Algebraic Circuit Complexity. *Theory Comput.*, 17:1–88, 2021.

[GGdOW20] Ankit Garg, Leonid Gurvits, Rafael Mendes de Oliveira, and Avi Wigderson. Operator Scaling: Theory and Applications. *Found. Comput. Math.*, 20(2):223–290, 2020.

[GP18] Joshua A. Grochow and Toniann Pitassi. Circuit Complexity, Proof Complexity, and Polynomial Identity Testing: The Ideal Proof System. *J. ACM*, 65(6):37:1–37:59, 2018.

[Gro26] Joshua A. Grochow. Polynomial Identity Testing and the Ideal Proof System: PIT Is in NP If and Only If IPS Can Be p-Simulated by a Cook-Reckhow Proof System. In *Proceedings of the 58th Annual ACM Symposium on Theory of Computing, STOC 2026, Salt Lake City, UT, USA, June 22-26, 2026*, pages 1569–1580. ACM, 2026.

[HK91] Tero Harju and Juhani Karhumäki. The Equivalence Problem of Multitape Finite Automata. *Theor. Comput. Sci.*, 78(2):347–355, 1991.

[HR25] Johan Håstad and Kilian Risse. On Bounded Depth Proofs for Tseitin Formulas on the Grid; Revisited. *SIAM Journal on Computing*, 54(5):FOCS22–288–FOCS22–339, 2025.

[IQS17] Gábor Ivanyos, Youming Qiao, and K. V. Subrahmanyam. Non-commutative Edmonds' problem and matrix semi-invariants. *Comput. Complex.*, 26(3):717–763, 2017.

[IQS18] Gábor Ivanyos, Youming Qiao, and K. V. Subrahmanyam. Constructive non-commutative rank computation is in deterministic polynomial time. *Comput. Complex.*, 27(4):561–593, 2018.

[Jeř25] Emil Jeřábek. A Simplified Lower Bound for Implicational Logic. *Bulletin of Symbolic Logic*, 31(1):53–87, 2025.

[KPW95] Jan Krajíček, Pavel Pudlák, and Alan R. Woods. An Exponential Lower Bound to the Size of Bounded Depth Frege Proofs of the Pigeonhole Principle. *Random Structures & Algorithms*, 7(1):15–39, 1995.

[KVV20] Igor Klep, Victor Vinnikov, and Jurij Volčič. Multipartite Rational Functions. *Documenta Math.*, 25:1285–1313, 2020.

[LTW18] Fu Li, Iddo Tzameret, and Zhengyu Wang. Characterizing Propositional Proofs as Noncommutative Formulas. *SIAM J. Comput.*, 47(4):1424–1462, 2018.

[Nis91] Noam Nisan. Lower Bounds for Non-Commutative Computation (Extended Abstract). In *Proceedings of the 23rd Annual ACM Symposium on Theory of Computing, May 5-8, 1991, New Orleans, Louisiana, USA*, pages 410–418, 1991.

[PBI93] Toniann Pitassi, Paul Beame, and Russell Impagliazzo. Exponential Lower Bounds for the Pigeonhole Principle. *Computational Complexity*, 3(2):97–140, 1993.

[PRST16] Toniann Pitassi, Benjamin Rossman, Rocco A. Servedio, and Li-Yang Tan. Polylogarithmic Frege Depth Lower Bounds via an Expander Switching Lemma. In *Proceedings of the 48th Annual ACM Symposium on Theory of Computing*, STOC '16, pages 644–657. ACM, 2016.

[PT16] Toniann Pitassi and Iddo Tzameret. Algebraic proof complexity: progress, frontiers and challenges. *ACM SIGLOG News*, 3(3):21–43, 2016.

[Raz13] Ran Raz. Tensor-Rank and Lower Bounds for Arithmetic Formulas. *J. ACM*, 60(6):40:1–40:15, 2013.

[Rec76] Robert A. Reckhow. *On the lengths of proofs in the propositional calculus*. Ph.d. thesis, University of Toronto, 1976. Technical Report #87.

[RS59] Michael O. Rabin and Dana S. Scott. Finite Automata and Their Decision Problems. *IBM J. Res. Dev.*, 3(2):114–125, 1959.

[RS05] Ran Raz and Amir Shpilka. Deterministic polynomial identity testing in non-commutative models. *Computational Complexity*, 14(1):1–19, 2005.

[Wor13] James Worrell. Revisiting the Equivalence Problem for Finite Multitape Automata. In *Automata, Languages, and Programming - 40th International Colloquium, ICALP 2013, Riga, Latvia, July 8-12, 2013, Proceedings, Part II*, pages 422–433, 2013.

[Zip79] R. Zippel. Probabilistic algorithms for sparse polynomials. In *Proc. of the Int. Sym. on Symbolic and Algebraic Computation*, pages 216–226, 1979.

A Partially Commutative Proof Systems

In this section we first define partially commutative polynomials and the partially commutative proof system ($\mathsf{PC}_{m,n}$-IPS for $m \leq n$). We then prove that it is sound and complete, discuss its verifiability and show that it simulates the *Frege* proof system.

Definition A.1 (Partially Commutative Polynomial)**.** *Let $X = X_1 \sqcup X_2$ be such that $|X| = n$ and let $\mathbb{F}$ be a field. A partially commutative monomial $\mathsf{mon} \in \mathbb{F}\langle X_1 \sqcup X_2 \rangle$ is uniquely determined by its restrictions to X_1 and to X_2: $\mathsf{mon}|_{X_1} = \mathsf{mon}_{X_1} \in \mathbb{F}\langle X_1 \rangle$ and $\mathsf{mon}|_{X_2} = \mathsf{mon}_{X_2} \in \mathbb{F}\langle X_2 \rangle$, where mon_{X_1} and mon_{X_2} are noncommutative monomials. All variables in X_1 commute with all variables in X_2. Hence, for $\mathsf{mon}, \mathsf{mon}' \in \mathbb{F}\langle X_1 \sqcup X_2 \rangle$, we have $\mathsf{mon} = \mathsf{mon}'$ if and only if $\mathsf{mon}|_{X_1} = \mathsf{mon}'|_{X_1}$ and $\mathsf{mon}|_{X_2} = \mathsf{mon}'|_{X_2}$. This induces an equivalence relation $\sim$ on the set of monomials.*[12]

A partially commutative polynomial $f \in \mathbb{F}\langle X_1 \sqcup X_2 \rangle$ is an $\mathbb{F}$-linear combination of the partially commutative monomials over X_1 and X_2.[13] *A partially commutative polynomial over p buckets, say $f \in \mathbb{F}\langle X_1 \sqcup \cdots \sqcup X_p \rangle$, can be defined analogously for any.* ◊

[12] For example, consider two monomials $\mathsf{mon}_1 = x_{1,1}x_{2,1}x_{1,2}x_{2,2}$ and $\mathsf{mon}_2 = x_{2,1}x_{1,1}x_{2,2}x_{1,2}$. These two monomials are the same with $\mathsf{mon}_1|_{X_1} = \mathsf{mon}_2|_{X_1} = x_{1,1}x_{1,2}$ and $\mathsf{mon}_1|_{X_2} = \mathsf{mon}_2|_{X_2} = x_{2,1}x_{2,2}$.

[13] It is sometimes useful to think of polynomials in $\mathbb{F}\langle X_1 \sqcup X_2 \rangle$ as a noncommutative polynomial in $\mathbb{F}\langle X_2 \rangle\langle X_1 \rangle$. That is, f can be uniquely written as $f = \sum_{m_i \in \mathbb{F}\langle X_1 \rangle} m_i \cdot f_i$ where $f_i \in \mathbb{F}\langle X_2 \rangle$.

Note that a commutative polynomial $f \in \mathbb{F}[X]$ can be thought of as a partially commutative polynomial in $\mathbb{F}\langle X_1, \ldots, X_n\rangle$ where $|X_i| = 1$. As mentioned in the introduction, inspired by this, we define a generalization of both IPS and NC-IPS. In particular, we define $\mathsf{PC}_{m,n}$-IPS where $\mathsf{PC}_{1,n}$-IPS $\equiv$ NC-IPS and $\mathsf{PC}_{n,n}$-IPS $\equiv$ IPS if the total number of variables is n.

We recall the definition here.

Definition 1.1. (Partially commutative IPS ($\mathsf{PC}_{m,n}$-IPS)) *Let $\mathbb{F}$ be a field and $X = X_1 \sqcup \cdots \sqcup X_m$ be a set of partially commutative variables*[14] *with $|X| = n$ and for every $i \in [m]$, $|X_i| = n_i$.*

Assume that $f_1 = f_2 = \cdots = f_\ell = 0$ is a set of partially commutative polynomial equations from $\mathbb{F}\langle X\rangle$ and suppose that the following equations (axioms) are included among the f_i's.

- ***Boolean axioms**: $x^2 - x$ for every $x \in X$*
- ***Commutator axioms**: $x_{i,j}x_{i,j'} - x_{i,j'}x_{i,j}$ for every $i \in [m]$, $j \neq j' \in [n_i]$.*

A partially commutative IPS proof ($\mathsf{PC}_{m,n}$-IPS) of unsatisfiability of the system $\{f_i\}$ is a polynomial $P(X, z_1, \ldots, z_\ell)$ represented by a partially commutative formula over the quotient of $\mathbb{F}\langle X, Z\rangle$ by the prescribed commuting relations among the X-variables only; in particular, no additional commuting relations involving the placeholder variables $Z = \{z_1, \ldots, z_\ell\}$ are imposed. We require

1. *$P(X, \overline{0}) = 0$;*
2. *$P(X, f_1, \ldots, f_\ell) = 1$,*

where substitution is the algebra homomorphism that fixes the X-variables and sends $z_i \mapsto f_i$. The size of a $\mathsf{PC}_{m,n}$-IPS proof is the minimum size of a partially commutative formula computing $P(X, Z)$. ◊

Remark A.2. *In Definition 1.1, if $P\langle X, Z\rangle$ is computed by a model $\mathcal{C}$ (can be a circuit, ABP or formula), we denote the proof system by $\mathcal{C}$-$\mathsf{PC}_{m,n}$-IPS. If $\mathcal{C}$ is circuits or formulas, then each product gate has designated left and right children; the leaves are labeled either by X variables or the set of placeholder variables Z with the convention specified in Definition 1.1. In the case of ABPs, the order of multiplication is maintained layer-wise. The size of a $\mathcal{C}$-$\mathsf{PC}_{m,n}$-IPS proof is the minimal size of a $\mathcal{C}$ (circuit or ABP or formula) that computes the proof polynomial $P(X, Z)$.* ◊

Remark A.3. *Analogous to the noncommutative setting, if we consider the partially commutative ring $\mathbb{F}\langle X_1 \sqcup X_2\rangle$ with $|X_1| = |X_2| = n$ and the set of equations $\{f_1 = 0, \ldots, f_m = 0 : f_i \in \mathbb{F}\langle X_1 \sqcup X_2\rangle\}$, a common solution is a set of $d \times d$ matrices (say $\overline{A} = \{A_1, \ldots, A_n\}$ and $\overline{B} = \{B_1, \ldots, B_n\}$) for some $d > 1$, such that the matrices from $\overline{A}$ and $\overline{B}$ commute with each other and that $f_i(\overline{A}, \overline{B}) = 0$.*

For the partially commutative ring $\mathbb{F}\langle X_1 \sqcup \cdots \sqcup X_m\rangle$ where $|X_i| = n_i$ for every $i \in [m]$, a set of equations $\{f_1 = 0, \ldots, f_\ell = 0\}$ have a common solution if there exists a set of $d \times d$ matrices (say $\overline{A_1} = \{A_{1,1}, \ldots, A_{1,n_1}\}, \ldots, \overline{A_m} = \{A_{m,1}, \ldots, A_{m,n_m}\}$) for some $d > 1$, such that the matrices from $\overline{A_i}$ and $\overline{A_j}$ commute with each other for every i, j and that $f_i(\overline{A_1}, \ldots, \overline{A_m}) = 0$. ◊

[14] For every i, the variables within X_i are noncommuting, but for every $i \neq j$, variables from X_i and X_j commute.

A.1 Soundness and Completeness

Lemma A.4. *For every m, n with $m \leq n$, the proof system* $\mathsf{PC}_{\mathrm{m,n}}$-$\mathsf{IPS}$ *is sound and complete for CNFs.*

Proof. Soundness follows from the definition of the system. Assume we have a polynomial $P(X, Z)$ that satisfies the properties $(1) : P(X, \overline{0}) = 0$ and $(2) : P(X, f_1, \ldots, f_\ell) = 1$. If the set of equations $\{f_i\}$ has a common solution $\overline{A} = (A_{1,1}, \ldots, A_{m,n_m})$, where each $A_{i,j} \in \mathbb{F}^{d\times d}$ for some $d \geq 1$, then we obtain a contradiction by substituting the variables with $\overline{A}$ since

$$(1) \implies P(\overline{A}, \overline{0}) = 0_{d\times d} \neq I_{d\times d} = P(\overline{A}, \overline{0}) \impliedby (2)$$

To prove *completeness*, we first claim the following.

Claim A.5. *Let $\{f_1 = 0, \ldots, f_\ell = 0 \ : \ f_i \in \mathbb{F}\langle X_1 \sqcup \cdots \sqcup X_m\rangle\}$ (with $|X_i| = n_i$ for each $i \in [m]$ and $\sum_{i\in[m]} |X_i| = n$) be a system of equations that includes the* boolean axioms *and the* commutator axioms. *Then the given system has a common solution in $\mathbb{F}^n$ if and only if there exist $d \geq 1$, set of matrices $\{A_{1,1}, \ldots, A_{m,n_m}\} \subseteq \mathbb{F}^{d\times d}$ which satisfy the given system.*

Proof. Suppose the given system has a common solution in $\mathbb{F}^n$, say $\overline{a} = (a_{1,1}, \ldots, a_{m,n_m})$. Taking $d = 1$ and $A_{ij} = a_{ij}$ for every $i \in [m], j \in [n_i]$, we trivially get the required statement.

For the converse direction, we give the full proof for $m = 1$ and observe that the proof follows analogously for larger m. Let $d \times d$ matrices $\overline{A} = (A_1, \ldots, A_n) \in (\mathbb{F}^{d\times d})^n$ be a common solution to the given system of equations.

Since the set of equations includes boolean axioms, $A_i^2 = A_i$ for every $i \in [n]$, the minimal polynomial of A_i must divide $x_i(x_i - 1)$. Thus the minimal polynomial has roots $0, 1$ with multiplicity 1, implying that each A_i is diagonalizable, with diagonal entries $0, 1$.

Moreover, the matrices satisfy the commutator axioms. That is, for every $i \neq j$, $A_iA_j = A_jA_i$. This implies that the matrices are simultaneously diagonalizable. That is, there exists an invertible matrix $R \in \mathbb{F}^{d\times d}$ such that for every i, $RA_iR^{-1} = D_i = \mathrm{Diag}(a_1^{(i)}, a_2^{(i)}, \ldots, a_d^{(i)})$ is a diagonal matrix with diagonal entries $a_j^{(i)} \in \{0, 1\}$. Observe that, for every $i \in [\ell]$,

$$f_i(A_1, \ldots, A_n) = 0 \implies R \cdot f_i(\overline{A}) \cdot R^{-1} = 0 \implies f_i(RA_1R^{-1}, \ldots, RA_nR^{-1}, \overline{a}) = 0.$$

Now for every monomial over X_1 in f_i, say $\mathsf{mon} = \prod_{j\in T} x_j \in f_i$ for some $T \subseteq [n]$,

$$R \cdot \mathsf{mon}(\overline{A}) \cdot R^{-1} = R(\prod_{j\in T} A_j)R^{-1} = \prod_{j\in T}(R \cdot A_j \cdot R^{-1}).$$

Thus, for every $i \in [\ell]$, $f_i(D_1, \ldots, D_n) = 0$ where each $D_j = \mathrm{Diag}(a_1^{(j)}, a_2^{(j)}, \ldots, a_d^{(j)})$ with $a_j^{(j)} \in \{0, 1\}$. Here, the left hand side is a $d \times d$ dimensional 0 matrix, with the (t, t)-th entry being $f_i(a_1^{(t)}, a_2^{(t)}, \ldots, a_n^{(t)})$. Since this is true for every f_i, we obtain a set of solutions over $\mathbb{F}$ to the system of equations, namely $\left\{(a_1^{(t)}, \ldots, a_n^{(t)}) \ : \ t \in [d]\right\}$.

For $m > 1$, the proof follows analogously since the matrices corresponding to different buckets commute (by definition) and we have commutator axioms for variables within each bucket. Hence all the matrices from the solution set mutually commute, and this gives an invertible matrix R that simultaneously diagonalizes every matrix. This proves the claim. □

Now we use Claim A.5 to prove the completeness. Consider the system of unsatisfiable equations $\{f_1 = 0, \ldots, f_\ell = 0 \mid f_i \in \mathbb{F}\langle X_1 \sqcup \cdots \sqcup X_m\rangle\}$ that includes the *boolean and commutator* axioms. Here $|X_i| = n_i$ for every $i \in [m]$ and $\sum_{i\in[m]} n_i = n$. We want to demonstrate the existence of the proof polynomial $P(X, Z) \in \mathbb{F}\langle X, Z\rangle$.

Moreover, assume that the axioms of the system other than the boolean and commutator axioms are obtained by a suitable arithmetization of an unsatisfiable CNF (as described below).

$$T(x_{ij}) = (1 - x_{i,j}) \qquad T(\neg x_{ij}) = x_{i,j} \qquad T(A \vee B) = T(A) \cdot T(B).$$

For the partially commutative setting, we fix an order of variables in each bucket and multiply them accordingly.

Now, for every $t \in [m]$, let $\mathcal{I}_t$ be the two sided ideal generated by the commutators $\{x_{t,j}x_{t,j'} - x_{t,j'}x_{t,j} \mid j' \neq j \in [n_t]\}$ and let $\pi : \mathbb{F}\langle X_1 \sqcup \cdots \sqcup X_m\rangle \to \mathbb{F}[X_1 \sqcup \cdots \sqcup X_m]$ be the surjective homomorphism obtained by quotienting by the two-sided ideal $\mathcal{I}_1 + \cdots + \mathcal{I}_m$. Equivalently, π imposes all within-bucket commutator relations. We denote $\pi(f_i)$ by f_i' for every $i \in [ell]$.

By Claim A.5, we know that partially commutative unsatisfiability in matrix algebra over $\mathbb{F}$ implies commutative unsatisfiability over $\mathbb{F}$. Since the *boolean axioms* are included in $\{f_i\}$, every common zero over the algebraic closure $\overline{\mathbb{F}}$ must in fact be Boolean: each coordinate satisfies $z^2 - z = 0$, hence lies in $\{0, 1\} \subseteq \mathbb{F}$. Therefore a common zero over $\overline{\mathbb{F}}$ would already give a common zero over $\mathbb{F}$, contradicting unsatisfiability in matrix algebra (in particular, in dimension 1). Thus the commutative system has no common zero over $\overline{\mathbb{F}}$. So, by Hilbert's Nullstellensatz, there are polynomials $g_1', \ldots, g_\ell' \in \mathbb{F}[X_1 \sqcup \cdots \sqcup X_m]$ such that $\sum_{i=1}^{\ell} g_i' \cdot f_i' = 1$. Let $g_i \in \mathbb{F}\langle X_1 \sqcup \cdots \sqcup X_m\rangle$ be a polynomial such that $\pi(g_i) = g_i'$[15]. Then,

$$\pi\left(1 - \sum_{i\in[\ell]} g_i \cdot f_i\right) = 1 - \sum_{i\in[\ell]} \pi(g_i) \cdot \pi(f_i) = 1 - \sum_{i\in[\ell]} g_i' \cdot f_i' = 0.$$

But $(1 - \sum_{i\in[\ell]} g_i \cdot f_i) \in \mathcal{I}_1 + \cdots + \mathcal{I}_m$, and hence

$$1 - \sum_{i\in[\ell]} g_i \cdot f_i = \sum_{t=1}^{m} \sum_{i\in[\ell], j\in[n_t]} h_{t,i,j}(x_{t,i}x_{t,j} - x_{t,j}x_{t,i})\hat{h}_{t,i,j}$$

for suitable polynomials $h_{t,i,j}, \hat{h}_{t,i,j}$.

[15]Note that there can be multiple g_i such that $\pi(g_i) = g_i'$. We take one such g_i

For each commutator $x_{t,i}x_{t,j} - x_{t,j}x_{t,i}$, let $z_{t,i,j}$ denote the corresponding placeholder already present among Z (since the commutator axioms are among the f_i's). We can then define

$$P(X,Z) := \sum_{i\in[\ell]} g_i \cdot z_i + \sum_{t=1}^{m} \sum_{i\in[\ell],j\in[n_t]} h_{t,i,j} \cdot z_{t,i,j} \cdot \hat{h}_{t,i,j},$$

to be the proof polynomial since it clearly satisfies conditions (1) and (2) from Definition 1.1. □

A.2 Verifiability

From the definition of IPS, verifying a $\mathcal{C}$-IPS refutation amounts to solving the $\mathcal{C}$-PIT problem. For $\mathsf{PC}_{m,n}$-IPS, refutations can be verified in randomized polynomial time when $m = O(1)$, and in randomized quasi-polynomial time when $m = (\log s)^{O(1)}$, where s is the refutation size. This follows from the randomized identity testing algorithm for formulas (and more generally, ABPs) over $\mathbb{Q}\langle X_1, \ldots, X_m\rangle$ due to Worrell [Wor13].

Arvind, Chatterjee, and Mukhopadhyay [ACM24] recently gave a deterministic identity testing algorithm for formulas (and ABPs) over $\mathbb{Q}\langle X_1, \ldots, X_k\rangle$ for any constant k. This implies the following: for any constant k, $\mathsf{PC}_{k,n}$-IPS is a sound and complete refutation system for unsatisfiable propositional formulas (represented as partially commutative formulas), and refutations can be verified in deterministic polynomial time over $\mathbb{F} = \mathbb{Q}$.

Note that, the algorithm given in [ACM24] does not work over finite field in general as discussed in Subsection 1.3. Nevertheless, when $k = 2$, the algorithm of [ACM24] can be applied over an arbitrary field $\mathbb{F}$[16]. In fact the white-box deterministic polynomial time algorithm given in Theorem 1.3 can also be used for the verification of $\mathsf{PC}_{2,n}$-IPS proofs. Hence $\mathsf{PC}_{2,n}$-IPS refutations admit deterministic polynomial-time verification in this case.

A.3 $\mathsf{PC}_{m,n}$-IPS simulates the *Frege* proof system

The commutative IPS and the noncommutative IPS simulate the Frege proof system [GP18, LTW18] and since $\mathsf{PC}_{m,n}$-IPS (for $m \le n$) is a generalization of both, it should simulate Frege. We prove that this is indeed the case, for the sake of completeness.

We start with the standard translation between propositional formulas and algebraic formulas.

Definition A.6. *Given a propositional formula T defined over $\wedge, \vee, \neg$, and variables $X = \{x_1, \ldots, x_n\}$, we define*

$$\mathrm{Tr}'(x_i) := x_i, \qquad \mathrm{Tr}'(\mathit{false}) := 1, \qquad \mathrm{Tr}'(\mathit{true}) := 0$$

and by induction on the size of the formula,

$$\mathrm{Tr}'(\neg T_i) := 1 - \mathrm{Tr}'(T_i); \mathrm{Tr}'(T_1 \vee T_2) := \mathrm{Tr}'(T_1)\,\mathrm{Tr}'(T_2); \mathrm{Tr}'(T_1 \wedge T_2) := 1 - (1 - \mathrm{Tr}'(T_1))(1 - \mathrm{Tr}'(T_2)).$$

[16] Personal communication with the authors of [ACM24].

If T is a propositional tautology then the polynomial computed by $\mathrm{Tr}'(T)$ is 0 over every boolean evaluation, i.e. $\widehat{\mathrm{Tr}'(T)}(\bar{a}) = 0$ for every $\bar{a} \in \{0,1\}^n$. $\Diamond$

We now prove the following theorem.

Theorem A.7. *If a propositional tautology T has a size s* Frege *proof, then for every $m \geq 1$ with $m \leq n$, $\mathrm{Tr}'(T)$ has a $\mathsf{PC}_{m,n}$-IPS-proof of size* $\mathrm{poly}(s)$.

In particular, we show how $\mathsf{PC}_{m,n}$-IPS simulates Łukasiewicz calculus (Definition 2.3) in polynomial size for any $1 \leq m \leq n$.

Lemma A.8. ($\mathsf{PC}_{m,n}$-IPS simulates Łukasiewicz calculus-Frege: analogue of [LTW18, Lemma 3.3]) *Let $\mathbb{F}$ be a field. Further, let $X = X_1 \sqcup \cdots \sqcup X_m$ be partially commutative variables where $|X_i| = n_i$ for every $i \in m$ and $\sum_{i\in[m]} n_i = n$. Here the variables inside each X_i are noncommuting and variables from different buckets commute.*

Let π be a tree-like *Frege (Łukasiewicz calculus) proof of a propositional formula T, defined over X, from assumptions $\{F_1, \ldots, F_\ell\}$, and let the proof-lines be $\mathsf{line}_1, \mathsf{line}_2, \ldots, \mathsf{line}_s$. Let $\mathrm{Tr}'(\cdot)$ be the algebraic translation from Boolean formulas to $\mathbb{F}\langle X\rangle$ as in Definition A.6. For each $i \in [N]$ define the partially commutative algebraic translation $L_i := \mathrm{Tr}'(\mathsf{line}_i) \in \mathbb{F}\langle X\rangle$ and let $F := \big(\mathrm{Tr}'(F_1), \ldots, \mathrm{Tr}'(F_\ell)\big)$.*

*Finally, let B be the set of all **Boolean axioms**: $x(x-1)$ for every $x \in X$ and let C be the set of all **partial commutator axioms**: $xy - yx$ for every $x, y \in X_i$, $i \in [m]$. Then for every $i \in [s]$, there exists a partially commutative algebraic formula $\Phi_i(X, Y, Z, W) \in \mathbb{F}\langle X, Y, Z, W\rangle$, where $Y = (y_1, \ldots, y_\ell)$ and Z, W are placeholder vectors for the Boolean and commutator axioms, such that:*

1. $\Phi_i(X, \bar{0}, \bar{0}, \bar{0}) = 0$.
2. $\Phi_i(X, F, B, C) = L_i$.
3. $|\Phi_i| \leq \left(\sum_{j\in A_i} |L_j|\right)^4$, *where $A_i \subseteq [s]$ is the set of Frege proof-lines involved in deriving line_i in the tree-like proof π (i.e., the indices of the sub-tree rooted at line_i).*

In particular, Φ_s is a $\mathsf{PC}_{m,n}$-IPS proof of $\mathrm{Tr}'(T)$ from assumptions $\{\mathrm{Tr}'(F_1), \ldots, \mathrm{Tr}'(F_\ell)\}$, and its size is $\mathrm{poly}(|\pi|)$. Here by $|.|$ we denote the size of the formula computing the polynomial.

Proof. We follow the inductive construction of Lemma 3.3 in [LTW18]. The only modification is that whenever [LTW18] uses commutator axioms to swap adjacent variables, in the partially commutative setting it suffices to use commutators *within the same bucket*, since variables from different buckets commute by definition. Fix the (tree-like) proof π and its sequence of lines $\mathsf{line}_1, \ldots, \mathsf{line}_s$. As in [LTW18], we build Φ_i by induction on the derivation of line_i.

Placeholders. The formula $\Phi_i(X, Y, Z, W)$ is a partially commutative formula in the variables X and in placeholder variables Y, Z, W, where:

- Y will be substituted by the translated assumptions $\mathrm{Tr}'(F_j)$;

- Z will be substituted by the Boolean axioms $x(x-1)$;
- W will be substituted by the commutator axioms $xy - yx$ for x, y in the same bucket.

Condition (1) says Φ_i vanishes when all placeholders are set to 0, and (2) says that after substitution we obtain exactly L_i.

Base cases. **(B1) line_i is an assumption.** Suppose $\mathsf{line}_i = F_j$ for some $j \in [\ell]$. Then $L_i = \mathrm{Tr}'(F_j)$. Define

$$\Phi_i(X, Y, Z, W) \;:=\; y_j.$$

Clearly, $\Phi_i(X, \overline{0}, \overline{0}, \overline{0}) = 0$ and $\Phi_i(X, F, B, C) = \mathrm{Tr}'(F_j) = L_i$.

(B2) line_i is a Łukasiewicz calculus axiom instance. We treat explicitly the axiom scheme $A \to (B \to A)$. Note that $A \to (B \to A) \equiv \neg A \vee (\neg B \vee A)$, and hence by the definition of $\mathrm{Tr}'(\cdot)$,

$$L_i = \mathrm{Tr}'\big(\neg A \vee (\neg B \vee A)\big) = \mathrm{Tr}'(\neg A) \cdot \mathrm{Tr}'(\neg B \vee A) = (1-a) \cdot \big((1-b) \cdot a\big),$$

where $a := \mathrm{Tr}'(A)$ and $b := \mathrm{Tr}'(B)$. Thus,

$$L_i = (1-a) \cdot (1-b) \cdot a.$$

By [LTW18, Lemma 3.6], there exists a partially commutative formula $\Psi_a(X, Z)$ such that

$$\Psi_a(X, \overline{0}) = 0, \qquad \Psi_a(X, B) = (1-a) \cdot a, \qquad |\Psi_a| \le |a|^2.$$

Multiplying on the left by $(1-b)$ yields a certificate for

$$(1-b) \cdot (1-a) \cdot a.$$

Define

$$\Theta(X, Z) := (1-b) \cdot \Psi_a(X, Z).$$

Then $\Theta(X, \overline{0}) = 0$ and $\Theta(X, B) = (1-b) \cdot (1-a) \cdot a$.

Now swap the first two factors to obtain $(1-a) \cdot (1-b) \cdot a$. Using the analogue of Lemma 3.4 from [LTW18] with only intra-bucket commutator axioms, there exists a partially commutative formula $\varphi_{M,N}(X, W)$ such that

$$\varphi_{M,N}(X, \overline{0}) = 0, \qquad \varphi_{M,N}(X, C) = M - N, \qquad |\varphi_{M,N}| \le |M|^2 |N|^2,$$

where

$$M := (1-b) \cdot (1-a) \cdot a, \qquad N := (1-a) \cdot (1-b) \cdot a = L_i.$$

Finally define

$$\Phi_i(X, Y, Z, W) := \Theta(X, Z) \; - \; \varphi_{M,N}(X, W).$$

Then $\Phi_i(X, \overline{0}, \overline{0}, \overline{0}) = 0$ and after substitution we obtain

$$\Phi_i(X, F, B, C) = M - (M - N) = N = L_i.$$

The other Łukasiewicz calculus axioms are handled analogously, as in [LTW18].

Induction step: modus ponens. Assume line_i is derived by modus ponens from $\mathsf{line}_j = A$ and $\mathsf{line}_h = A \to B$, with $j, h < i$. By the definition of $\mathrm{Tr}'(\cdot)$, since $A \to B \equiv \neg A \vee B$, we have

$$L_j = \mathrm{Tr}'(A) = a, \qquad L_h = \mathrm{Tr}'(A \to B) = \mathrm{Tr}'(\neg A \vee B) = (1 - a) \cdot \mathrm{Tr}'(B) = (1 - a) \cdot c,$$

where $c := \mathrm{Tr}'(B)$. The conclusion is $\mathsf{line}_i = B$, hence $L_i = c$.

By induction, we have Φ_j, Φ_h satisfying (2)-(3) in Definition 1.1 . Define

$$\Phi_i(X, Y, Z, W) \; := \; \Phi_j(X, Y, Z, W) \cdot \mathrm{Tr}'(B) \; + \; \Phi_h(X, Y, Z, W).$$

Then $\Phi_i(X, \overline{0}, \overline{0}, \overline{0}) = 0$ and after substituting $(Y, Z, W) = (F, B, C)$ we obtain

$$\begin{aligned} \Phi_i(X, F, B, C) &= \mathrm{Tr}'(A) \cdot \mathrm{Tr}'(B) \; + \; (1 - \mathrm{Tr}'(A)) \cdot \mathrm{Tr}'(B) \\ &= \big(\mathrm{Tr}'(A) + (1 - \mathrm{Tr}'(A))\big) \cdot \mathrm{Tr}'(B) \\ &= 1 \cdot \mathrm{Tr}'(B) = \mathrm{Tr}'(B) = L_i. \end{aligned}$$

The size bound follows exactly as in the proof pf [LTW18, Lemma 3.3]. Since the present construction changes only which commutator axioms are available, the size analysis is unchanged.

Conclusion. Applying the construction to $i = s$ gives a partially commutative IPS proof of $L_s = \mathrm{Tr}'(T)$ from assumptions $\mathrm{Tr}'(F_1), \ldots, \mathrm{Tr}'(F_\ell)$ and the Boolean/partial-commutator axioms, with size $\mathrm{poly}(|\pi|)$. □

B Proof of Theorem 3.3

Throughout this section, let $X = X_1 \sqcup X_2, \qquad |X_1| = n_1, \qquad |X_2| = n_2, \qquad N = n_1 + n_2$. Thus, when we refer to the corresponding partially commutative proof system, it is $\mathsf{PC}_{2,N}$- IPS. Variables within each X_i are noncommuting, while variables from X_1 commute with variables from X_2. Throughout the Frege simulation the underlying field is $\mathbb{F} = \mathbf{GF}(2)$.

B.1 The Reflection Principle

In order to prove Theorem 3.3, we use the reflection principle.

Lemma B.1. ([LTW18, Lemma 4.3]) *There is a polynomial-size* Frege *proof of* $\neg\Phi$ *assuming that* $\neg\widetilde{F}(X,\overline{0})$ *and* $\widetilde{F}(X,\widetilde{P_\Phi})$ *have polynomial-size* Frege *proofs.*

Even though Li, Tzameret, and Wang [LTW18] prove this statement for noncommutative IPS, the same proof applies in the partially commutative setting. Once Frege proofs of $\neg\widetilde{F}(X,\overline{0})$ and $\widetilde{F}(X,\widetilde{P_\Phi})$ are available, the remainder of the reflection argument is purely propositional.

Therefore, our goal is to obtain Frege proofs of the two tautologies in Equation 3.2. In particular, using Lemma B.1, Theorem 3.3 follows directly from the following theorem.

Theorem B.2. *Let* $X = X_1 \sqcup X_2$ *be partially commutative variables, where* $|X_1| = n_1$, $|X_2| = n_2$, *and* $N = n_1 + n_2$. *Let* F *be a partially commutative formula of size* s *over* $\mathbb{F}$, *with variable set* X, *that computes the zero polynomial. Then there exists a constant* $c > 0$ *such that the Boolean formula* $\neg\widetilde{F}(X)$ *has a* Frege *proof of size* $s^{(\log s)^c}$.

The remainder of the section is dedicated to the proof of Theorem B.2. The remainder of the section is dedicated to the proof of Theorem B.2.

Following the notation of [LTW18], for Boolean formulas $\widetilde{F}$ and $\widetilde{G}$, we write $\widetilde{F} \vdash \widetilde{G}$ to denote that there exists a Frege proof of $\widetilde{G}$ from the axiom $\widetilde{F}$ of size polynomial in $|\widetilde{G}|$. That is, $\widetilde{G}$ can be derived from $\widetilde{F}$ by purely syntactic manipulations using standard Frege inference rules and Boolean equivalences, including associativity and distributivity, as well as identities such as $\text{False} \oplus \widetilde{F} \equiv \widetilde{F}$, $\text{True} \vee \widetilde{F} \equiv \text{True}$, $\text{True} \oplus \widetilde{F} \equiv \neg\widetilde{F}$, $\text{False} \wedge \widetilde{F} \equiv \text{False}$, $\text{True} \wedge \widetilde{F} \equiv \widetilde{F}$, and similar Boolean equivalences.

Similarly, for two vectors of formulas $\overline{F}$ and $\overline{G}$, we write $\overline{F} \vdash \overline{G}$ to denote that, for each index, the corresponding entry of $\overline{G}$ can be derived from the corresponding entry of $\overline{F}$ by a Frege proof of size polynomial in the size of that entry.

B.2 Homogenization of Partially Commutative Formulas

Note that the results given in this section are true for any partially commutative formula with a k-partition of variables for any constant k. Let F be a partially commutative formula over $\mathbb{F}$ with the variable set $X = X_1 \sqcup \ldots \sqcup X_k$.

Given a partially commutative monomial $m = \prod_{i\in[k]} m_i$ with m_i being noncommutative in X_i for every $i \in [k]$, define the degree signature of m to be $d(m) := (d_1, \ldots, d_k)$ if the degree of m_i is d_i for every $i \in [k]$. A partially commutative polynomial f is homogeneous if every monomial of f has the same degree signature. Thus any partially commutative polynomial computed by a

formula F, say $\hat{F}$, can be uniquely written as

$$\hat{F} = \sum_{\substack{d_1,\ldots,d_k \le d \\ \sum d_i \le d}} \hat{F}_{(d_1,\ldots,d_k)}$$

where the degree of $\hat{F}$ is d and $\hat{F}_{(d_1,\ldots,d_k)}$ is the homogeneous component of $\hat{F}$ with degree signature $(d_1,\ldots,d_k)$.

Lemma B.3. *Let F be a partially commutative formula of size s and depth $O(\log s)$ computing a polynomial in $\mathbb{F}\langle X_1 \sqcup \ldots \sqcup X_k\rangle$. Moreover, let $F_{(d_1,\ldots,d_k)}$ be the homogeneous component of the polynomial computed by F with degree signature $(d_1,\ldots,d_k)$. Then there exists a $s^{O(k\log s)}$ size Frege proof of*

$$\widetilde{F}(X) \longleftrightarrow \bigoplus_{(d_1,\ldots,d_k)} \widetilde{F}_{(d_1,\ldots,d_k)}. \tag{B.4}$$

This is analogous to [LTW18, Lemma 4.6] and the proof follows along the same lines.

Proof. We first show the homogenization of [Raz13] in the partially commutative setting; then, we show how *Frege* can prove it.

Raz's Homogenization in the Partially Commutative Setting Let $X = X_1 \sqcup X_2 \sqcup \cdots \sqcup X_k$ be the variable set. Since the size of the formula F is s, the degree of the polynomial computed by F is at most $s+1$. Furthermore, the individual degree from each bucket can be at most $s+1$. Given any node $u \in F$, denote the product depth of u by δ_u. We assume the depth of F is $O(\log s)$ and hence, $\delta_u \le O(\log s)$ for any $u \in F$. The idea is to track the increment of any monomial's degree signature along the path from $u \to r$ with r being the root. This is done by considering all monotone non increasing functions $N_u := \left\{ D^u : \{0,1,\ldots,\delta_u\} \to \{0,1,\ldots,s+1\}^k \right\}$. The number of such functions is $|N_u| \le \binom{s+\delta_u+2}{\delta_u+1}^k \le s^{O(k\log s)}$. So, first we construct the homogeneous formula F' with $|N_u|$ many nodes for every node $u \in F$. We identify 0 with the root r in the function $D^u \in N_u$. For $(i_1,\ldots,i_k) \in [0,1,\ldots,s+1]^k$, $D^u_{i_1,\ldots,i_k}$ denotes all the functions $D^u \in N_u$ such that $D^u(\delta_u) = (i_1,\ldots,i_k)$. We want to prove for every node $u \in F$,

$$\left(\bigoplus_{(d_1,\ldots,d_k)} \widetilde{F'}_{(u,D^u_{d_1,\ldots,d_k})} \right) \longleftrightarrow \widetilde{F}_u. \tag{B.5}$$

If we prove Equation B.5 for the root r, we prove Equation B.4 within Frege.

Construction of F' and Frege proof If $u \in F$ is a leaf, then for every $D^u \in N_u$, the node (u, D^u) is a leaf. If u is labeled by a field element and $D^u(\delta_u) = \underbrace{(0,\ldots,0)}_{k}$, then (u, D^u) is labeled by the same field element, and if $D^u(\delta_u) \neq (0,\ldots,0)$ then (u, D^u) is labeled by 0. If u is labeled from X_1 (respectively X_i for $i \in [k]$), then (u, D^u) is labeled by the same variable only if $D^u(\delta_u) =$

$(1,0,\ldots,0)$ (respectively $D^u(\delta_u) = (0,\ldots,1,\ldots,0)$). Otherwise, we label it by 0. In this case, we can clearly prove Equation B.5 within Frege.

Inductive Case : Let $u,v,w \in F$ such that $u = v + w$. For every $D^u \in N_u$, let $D^v \in N_v$ be the function that agrees with D^u on $\{0,1,\ldots,\delta_u\}$ and $D^v(\delta_v) = D^u(\delta_u)$. Similarly, consider such $D^w \in N_w$. Then we have the following connection $(u,D^u) = (v,D^v) + (w,D^w)$ in F'. This implies

$$\begin{aligned} \widehat{F'}_{(u,D^u)} &:= \widehat{F'}_{(v,D^v)} + \widehat{F'}_{(w,D^w)} \\ \text{If } D^u(\delta_u) = (i_1,\ldots,i_k) &\text{ then we have the following,} \\ \widehat{F'}_{(u,D^u_{i_1,\ldots,i_k})} &:= \widehat{F'}_{(v,D^v_{i_1,\ldots,i_k})} + \widehat{F'}_{(w,D^w_{i_1,\ldots,i_k})}. \end{aligned} \tag{B.6}$$

Hence, we have the following tautology,

$$\widetilde{F'}_{(u,D^u_{i_1,\ldots,i_k})} \longleftrightarrow \widetilde{F'}_{(v,D^v_{i_1,\ldots,i_k})} \bigoplus \widetilde{F'}_{(w,D^w_{i_1,\ldots,i_k})}. \tag{B.7}$$

By the inductive hypothesis,

$$\begin{aligned} & \widetilde{F}_v \leftrightarrow \bigoplus_{(i_1,\ldots,i_k)} \widetilde{F'}_{(v,D^v_{i_1,\ldots,i_k})} \qquad \widetilde{F}_w \leftrightarrow \bigoplus_{(i_1,\ldots,i_k)} \widetilde{F'}_{(w,D^w_{i_1,\ldots,i_k})} \\ \implies & \widetilde{F}_v \oplus \widetilde{F}_w \leftrightarrow \bigoplus_{(i_1,\ldots,i_k)} (\widetilde{F'}_{(v,D^v_{i_1,\ldots,i_k})} \oplus \widetilde{F'}_{(w,D^w_{i_1,\ldots,i_k})}) \qquad \text{we can prove} \\ \implies & \widetilde{F}_u \longleftrightarrow \bigoplus_{(i_1,\ldots,i_k)} \widetilde{F}_{(u,D^u_{i_1,\ldots,i_k})} \qquad \text{this follows from Equation B.7.} \end{aligned} \tag{B.8}$$

Let $u = v \times w$ and $D^u \in N_u$ be such that $D^u(u) = (d_1,\ldots,d_k)$.
Then, $(u,D^u) = \sum_{i_1=0}^{d_1}\sum_{i_2=0}^{d_2}\cdots\sum_{i_k=0}^{d_k}(v,D^v_{i_1,\ldots,i_k})(w,D^w_{d_1-i_1,\ldots,d_k-i_k})$. That is,

$$\begin{aligned} \widehat{F'}_{(u,D^u_{i_1,\ldots,i_k})} &:= \sum_{j_1=0}^{i_1}\sum_{j_2=0}^{i_2}\cdots\sum_{j_k=0}^{i_k} \widehat{F'}_{(v,D^v_{j_1,\ldots,j_k})}\widehat{F'}_{(w,D^w_{i_1-j_1,\ldots,i_k-j_k})} \\ \implies \widetilde{F'}_{(u,D^u_{i_1,\ldots,i_k})} &\leftrightarrow \oplus_{j_1=0}^{i_1}(\cdots(\oplus_{j_k=0}^{i_k}(\widetilde{F'}_{(v,D^v_{j_1,\ldots,j_k})} \wedge \widetilde{F'}_{(w,D^w_{i_1-j_1,\ldots,i_k-j_k})})\cdots). \end{aligned} \tag{B.9}$$

$$\begin{aligned} \widetilde{F}_u \longleftrightarrow \widetilde{F}_v \wedge \widetilde{F}_w &\leftrightarrow \Big(\bigoplus_{\substack{(i_1,\ldots,i_k)\\ 0\le i_1,\ldots,i_k\le s_v+1}} \widetilde{F'}_{(v,D^v_{i_1,\ldots,i_k})}\Big) \wedge \Big(\bigoplus_{\substack{(j_1,\ldots,j_k)\\ 0\le j_1,\ldots,j_k\le s_w+1}} \widetilde{F'}_{(w,D^w_{j_1,\ldots,j_k})}\Big) \\ &\leftrightarrow \bigoplus_{\substack{(j_1,\ldots,j_k)\\ 0\le j_1,\ldots,j_k\le s_u+1}} (\oplus_{j_1=0}^{i_1}\cdots(\oplus_{j_k=0}^{i_k}(\widetilde{F'}_{(v,D^v_{j_1,\ldots,j_k})} \wedge \widetilde{F'}_{(w,D^w_{i_1-j_1,\ldots,i_k-j_k})}))\cdots) \\ &\leftrightarrow \bigoplus_{\substack{(i_1,\ldots,i_k)\\ 0\le i_1,\ldots,i_k\le s_u+1}} \widetilde{F'}_{(u,D^u_{i_1,\ldots,i_k})} \qquad \text{follows from Equation B.9.} \end{aligned} \tag{B.10}$$

To show within Frege that the homogeneous partially commutative formula F is identically 0, we first convert F into a layered homogeneous ABP A_F that computes the same polynomial. Following [LTW18], we establish equations that witness A_F computes the zero polynomial.

Since Frege operates on formulas rather than ABPs, we translate this argument back to formulas by associating each partial ABP in A_F with a unique induced formula in F that computes the same polynomial. Let $\widetilde{F}$ be the Boolean formula obtained from the homogeneous partially commutative formula F over $\mathbf{GF}(2)$ via Definition 3.1.

B.3 Translation Between Formulas and ABPs

We first eliminate Boolean constants from $\widetilde{F}$ using the following lemma from [LTW18]. That is, we transform $\widetilde{F}$ into a constant-free formula by applying standard Boolean equivalences. A Formula is called constant-free if there is no leaf labelled by constants. Although the lemma is stated for noncommutative formulas in [LTW18], it applies verbatim to partially commutative formulas.

Lemma B.11. ([LTW18, Lemma 4.10]) *Let F be a non-constant partially commutative homogeneous formula over $\mathbf{GF}(2)$ that computes the zero polynomial. Then there exists a constant-free partially commutative formula F' of size $\mathrm{poly}(|F|)$ that computes the zero polynomial, such that $\widetilde{F} \vdash \widetilde{F'}$.*

Next, given a partially commutative homogeneous constant-free formula F, we construct the corresponding homogeneous partially commutative ABP A_F following the construction of [Nis91]. The construction proceeds inductively as follows.

- If $u = v + w$ is a node of F and the ABPs A_v and A_w have already been constructed, then A_u is obtained by taking the parallel composition of A_v and A_w, merging their source nodes into a single source and their sink nodes into a single sink. The resulting ABP computes the sum of the polynomials computed by A_v and A_w.
- If $u = v \times w$, then A_u is obtained by sequentially composing A_v and A_w: the sink of A_v is identified with the source of A_w, and the source of A_v and the sink of A_w serve as the source and sink of A_u, respectively.

Tracking computation in F corresponding to computation done A_F

The arguments are the same as those of [LTW18], but we recall them here as well for completeness. Firstly, we need the following two notions defined in their work.

Definition B.12 (Induced formula [LTW18])**.** *Given a formula F, let F' be a sub-formula of F. Assume $\{u_1, \ldots, u_s\}$ are nodes in F' and $\{\alpha_1, \ldots, \alpha_s\}$ are elements of $\mathbb{F}$. Then $F'(u_1 = \alpha_1, \ldots, u_s = \alpha_s)$ is called the induced formula from F.* ◊

Definition B.13 (v-part [LTW18])**.** *Let F be a formula and A_F be the corresponding ABP for F with source and sink nodes s, t, respectively. For any node $v \in A_F$, the v-part of F is an induced formula computing the*

polynomial $A_F[v,t]$. Although a formula may admit multiple v-parts in general, for homogeneous formulas and their corresponding homogeneous ABPs, we define a canonical v-part for every node $v \in A_F$, denoted by F_v^*. ◊

We now consider a map $g : \{\text{nodes of } A_F\} \to \{\text{nodes of } F\} \cup \{\emptyset\}$ where $g(u) = v$ if the polynomial computed by F at v is the polynomial computed between u and t in A_F.

The definition is given by induction and can be found in [LTW18]. We recall it here.

Base case. Let F be the formula $u = x$. The corresponding ABP A_F consists of a single edge $s \xrightarrow{x} t$. We define $g(s) := u$ and $g(t) := \emptyset$.

Inductive case (addition). Let $F = G + H$, where u, v, w are the roots of F, G, H, respectively, with $u = v + w$. Let A_G, A_H be the corresponding ABPs, and let g_1, g_2 be the associated maps with $g_1(s) = v$ and $g_2(s) = w$. We abuse notation by denoting the source and sink nodes of both ABPs by s and t.

The ABP A_F is obtained using the parallel composition of A_G and A_H. The map g is defined by

$$g(u_i) = \begin{cases} u, & \text{if } u_i \text{ is the source node,} \\ g_1(u_i), & \text{if } u_i \in A_G, \\ g_2(u_i), & \text{if } u_i \in A_H, \\ \emptyset, & \text{if } u_i \text{ is the sink node.} \end{cases} \tag{B.14}$$

Inductive case (multiplication). If $F = G \cdot H$, then A_F is obtained by sequentially composing A_G and A_H. The source of A_F is the source of A_G, and the sink of A_G is identified with the source of A_H. The map g is defined as the union of g_1 and g_2, except at the merged node u_t, where we set

$$g(u_t) := g_2(s),$$

with s denoting the source of A_H.

Now, to derive the induced formulas explicitly corresponding to the nodes of the ABP, we need the following function.

Definition of the function D. We then define a function D that takes as input a homogeneous sub-formula F_u and a node $v \in F$, and outputs an induced sub-formula of F. Although D is defined for all nodes of F, we only apply it to nodes in the image of g. Note that constants introduced in D do not appear in the original constant-free formula F.

For every node $u \in F$, define $D(F_u, u) := F_u$. If u is not a node of F, then $D(F, u)$ is undefined. The function D is defined inductively from the leaves upwards.

Let G, H be homogeneous sub-formulas of F. For a $+$-node,

$$D(G+H,u) = \begin{cases} D(G,u)+0, & u \in G, \\ 0+D(H,u), & u \in H. \end{cases} \tag{B.15}$$

For a $\times$-node,

$$D(G\times H,u) = \begin{cases} D(G,u)\times H, & u \in G, \\ 1\times D(H,u), & u \in H. \end{cases} \tag{B.16}$$

Definition of the v-part. Finally, for a node $v \in A_F$, define $F_v^* := D(F, g(v))$. Let r be the root of F. Since $g(s) = r$, where s is the source node of A_F, we have $F_s^* = D(F,r) = F$, which computes the same polynomial $\hat{F}$ as the original homogeneous formula F.

Following the notation that has been defined, [LTW18] shows,

$$F_v^* \vdash \sum_{u\colon u \text{ has an incoming edge from } v} A_F[v,u]\cdot F_u^*.$$

That is, we can prove the following tautology within Frege:

$$F_v^* \quad \leftrightarrow \sum_{u\colon u \text{ has an incoming edge from } v} A_F[v,u]\cdot F_u^*.$$

Thus, it is enough to argue about the partially commutative ABP A_F, which is what we do now.

B.4 Refined Translation Between Formulas and ABPs

To handle partial commutativity, we represent polynomials as sums of monomials over a fixed bucket, with coefficients given by polynomials over the remaining buckets. We need this viewpoint to be incorporated directly into the underlying computational model (ABPs or formulas). Note that even though the following construction works for any constant k, we present it for $k = 2$.

In particular, starting from the homogeneous partially commutative ABP A_F over the variable set $X = X_1 \sqcup X_2$, we first *partially homogenize* it with respect to the first bucket X_1. The goal is to transform A_F into an equivalent ABP A'_F computing the same polynomial such that:

- each edge label of A'_F is a homogeneous linear form over X_1, and
- the coefficients of these linear forms are noncommutative ABPs over $\mathbb{F}\langle X_2\rangle$.

We achieve this by applying the homogenization procedure of Lemma B.17.

Lemma B.17. [ACM24, Lemma 12] *Let $f \in \mathbb{F}\langle X_1 \sqcup \ldots \sqcup X_k\rangle$ be a partially commutative polynomial of degree d computed by an ABP of size s. Then for any $1 \le i \le k$, we can efficiently homogenize the ABP*

over the variable set X_i, and the coefficients $\in \mathbb{F}\langle X_1 \sqcup \ldots X_{i-1} \sqcup X_{i+1} \sqcup \cdots \sqcup X_k\rangle$ can also be computed by partially commutative ABPs of size $O(sd)$.

B.4.1 Refined ABP Construction

Recall that F is a partially commutative homogeneous constant-free formula over the variable set $X = X_1 \sqcup X_2$, and A_F is the corresponding homogeneous partially commutative ABP of size s and depth d computing the same polynomial.

Thus, every edge label in A_F from $u_{i,j}$ to $u_{i+1,\ell}$ is of the form $L_{j,\ell} = L_{j,\ell}^{(1)}(X_1) + L_{j,\ell}^{(2)}(X_2)$, where each $L_{j,\ell}^{(i)}$ is a homogeneous noncommutative linear form over the variable set X_i.

First stage: degree tracking. We first construct an intermediate ABP $\widetilde{A}_F$ that explicitly tracks the accumulated X_1-degree. This is the first step in the homogenization procedure of [ACM24, Lemma 12].

For each node $u_{i,j}$ of A_F, introduce nodes $u_{i,j}^{(0)}, u_{i,j}^{(1)}, \ldots, u_{i,j}^{(d)}$ in $\widetilde{A}_F$, where the superscript records the accumulated X_1-degree.

Base Case: Let $s \xrightarrow{L_i} u_{1,i}$ in A_F, where $L_i = L_i^{(1)}(X_1) + L_i^{(2)}(X_2)$. Then $\widetilde{A}_F$ contains $s \xrightarrow{L_i^{(1)}} u_{1,i}^{(1)}$ and $s \xrightarrow{L_i^{(2)}} u_{1,i}^{(0)}$.

Inductive Step: If $u_{i,j} \xrightarrow{L_{j,\ell}} u_{i+1,\ell}$ is an edge of A_F, then for every $p \in \{0, \ldots, d\}$, $\widetilde{A}_F$ contains $u_{i,j}^{(p)} \xrightarrow{L_{j,\ell}^{(1)}} u_{i+1,\ell}^{(p+1)}$ when $p < d$, and $u_{i,j}^{(p)} \xrightarrow{L_{j,\ell}^{(2)}} u_{i+1,\ell}^{(p)}$.

Thus, an edge labeled by a linear form in X_1 increases the tracked degree by one, whereas an edge labeled by a linear form in X_2 preserves it.

For a node $u_{i,j}$ of A_F, we write $(A_F[s, u_{i,j}])^{(p)}$ to denote the X_1-degree-p homogeneous component of the polynomial computed by $A_F[s, u_{i,j}]$. The following claim follows directly from the construction.

Claim B.18. *For every node $u_{i,j}$ of A_F and every $p \in \{0, \ldots, d\}$, $\widetilde{A}_F[s, u_{i,j}^{(p)}] = \left(A_F[s, u_{i,j}]\right)^{(p)}$, where $(\cdot)^{(p)}$ denotes the X_1-degree-p homogeneous component.*

In particular, if t is the sink of A_F, then $\widetilde{A}_F[s, t^{(p)}]$ computes the X_1-degree-p homogeneous component of $A_F[s, t]$. The size of $\widetilde{A}_F$ is $O(sd)$.

Second stage: compression of the X_2-edges. The ABP $\widetilde{A}_F$ is not yet in the form required by the Witness Theorem: it contains edges that preserve the X_1-degree. Following the second step of the homogenization procedure in [ACM24, Lemma 12], we eliminate these degree-preserving edges.

For nodes $v^{(q)}$ and $u^{(q)}$ of $\widetilde{A}_F$ having the same X_1-degree q, let $C_{v,u}^{(q)} \in \mathbb{F}\langle X_2\rangle$ denote the sum of the weights of all paths from $v^{(q)}$ to $u^{(q)}$ that use only degree-preserving edges. Equivalently,

these are precisely the paths whose edge labels involve only variables from X_2. Since the degree-q subgraph is an acyclic subgraph of $\widetilde{A}_F$, every $C^{(q)}_{v,u}$ is computed by a noncommutative ABP over X_2 of size $O(sd)$.

Fix an X_1-degree $r \in \{1, \dots, d\}$. We construct an ABP $A'_{F,r}$ computing $\widetilde{A}_F[s, t^{(r)}]$. Its layers are indexed by the accumulated X_1-degree $q = 0, 1, \dots, r$.

The source is s, which lies in degree layer 0. For $1 \le q < r$, the nodes in degree layer q are the nodes $v^{(q)}$ which can occur immediately after the q-th X_1-edge of a path in $\widetilde{A}_F$. The final layer consists only of the sink $t^{(r)}$.

Consider $0 \le q < r-1$. For nodes $v^{(q)}$ and $w^{(q+1)}$, define the edge label $G^{(q)}_{v,w}$ by summing over all X_1-edges $u^{(q)} \xrightarrow{L^{(1)}_{u,w}} w^{(q+1)}$ of $\widetilde{A}_F$: $G^{(q)}_{v,w} := \sum_u C^{(q)}_{v,u} L^{(1)}_{u,w}$.

Since variables from X_1 commute with variables from X_2 in the partially commutative algebra, every such label can be written uniquely in the form $G^{(q)}_{v,w} = \sum_{j=1}^{n_1} x_{1,j} f^{(q)}_{v,w,j}$, where every $f^{(q)}_{v,w,j} \in \mathbb{F}\langle X_2 \rangle$ is computable by a noncommutative ABP of size $\mathrm{poly}(s, d)$.

For the final transition, we must also absorb all degree-preserving X_2-paths occurring after the last X_1-edge. For a node $v^{(r-1)}$, define $G^{(r-1)}_{v,t} := \sum_{u,w} C^{(r-1)}_{v,u} L^{(1)}_{u,w} C^{(r)}_{w,t}$. Again, using the commutation between the two buckets, $G^{(r-1)}_{v,t} = \sum_{j=1}^{n_1} x_{1,j} f^{(r-1)}_{v,t,j}$, where all the coefficients $f^{(r-1)}_{v,t,j}$ are computed by $\mathrm{poly}(s, d)$-size noncommutative ABPs over X_2.

Thus every edge of $A'_{F,r}$ goes from degree layer q to degree layer $q+1$, and every edge label has the form $\sum_{j=1}^{n_1} x_{1,j} f_j, \qquad f_j \in \mathbb{F}\langle X_2 \rangle$, with each f_j computed by a polynomial-size noncommutative ABP.

Every path in $\widetilde{A}_F$ from s to $t^{(r)}$ has a unique decomposition into r occurrences of X_1-edges, separated by possibly empty paths consisting only of X_2-edges. The construction above contracts precisely these intervening X_2-paths. Therefore, $A'_{F,r}[s, t^{(r)}] = \widetilde{A}_F[s, t^{(r)}] = (A_F[s, t])^{(r)}$.

Hence $A'_{F,r}$ computes exactly the X_1-degree-r homogeneous component of the polynomial computed by A_F. Moreover, $A'_{F,r}$ has polynomial size in s and d, and every coefficient occurring on an edge is computed by a $\mathrm{poly}(s, d)$-size noncommutative ABP over X_2.

For $r = 0$, the homogeneous component contains only variables from X_2. In this case no compression or witness construction is required: the resulting identity is a noncommutative identity over X_2 and can be handled directly using [LTW18].

B.4.2 Tracking the Partially Homogenized ABP by Induced Formulas

Let F be a fully homogeneous, constant-free, partially commutative formula over $\mathbb{F}\langle X_1 \sqcup X_2 \rangle$ obtained after Raz's homogenisation. Let A_F be the corresponding homogeneous layered ABP obtained from F using Nisan's construction.

By the construction of Li, Tzameret and Wang [LTW18], there exists a map $g : \mathrm{Nodes}(A_F) \to \mathrm{Nodes}(F)$ and a recursively defined operator D such that, for every node $v \in A_F$, $F^*_v := D(F, g(v))$ is an induced formula satisfying $\widehat{F^*_v} = A_F[v, t]$. Moreover, every formula F^*_v has size $s^{O(\log s)}$.

Fix an X_1-degree $r \in \{1, \ldots, d\}$ and consider the compressed ABP $A'_{F,r}$ constructed in the previous subsection. For simplicity, throughout this subsection we write A'_F for $A'_{F,r}$.

Recall that a node $v^{(q)}$ in the q-th layer of A'_F records that exactly q variables from X_1 have already been encountered along a path from the source. Therefore, from such a node to the sink $t^{(r)}$, the remaining X_1-degree is $r - q$.

The following lemma identifies the polynomial computed from every node of the compressed ABP to the sink.

Lemma B.19. *For every node $v^{(q)}$ of A'_F, where $0 \le q \le r$, $A'_F[v^{(q)}, t^{(r)}] = (A_F[v, t])^{(r-q)}$, where $(\cdot)^{(r-q)}$ denotes the homogeneous component of X_1-degree $r - q$.*

Proof. Consider first the intermediate degree-tracking ABP $\widetilde{A}_F$. A path from $v^{(q)}$ to $t^{(r)}$ increases the tracked X_1-degree from q to r. Consequently, it contains exactly $r - q$ edges labeled by linear forms in X_1. Hence, $\widetilde{A}_F[v^{(q)}, t^{(r)}] = (A_F[v, t])^{(r-q)}$.

The second compression does not change this path polynomial. Indeed, every path from $v^{(q)}$ to $t^{(r)}$ admits a unique decomposition into $r - q$ X_1-edges separated by paths involving only X_2-edges. The edge labels $G^{(q)}_{v,w}$ of A'_F are defined precisely by summing over these intervening X_2-paths. Therefore, $A'_F[v^{(q)}, t^{(r)}] = \widetilde{A}_F[v^{(q)}, t^{(r)}]$. Combining the two equalities proves the lemma. □

We now associate a formula with every node of A'_F. For every node $v \in A_F$, apply Raz's homogenization procedure to the induced formula F^*_v with respect to X_1. For a node $v^{(q)}$ of the compressed ABP A'_F, define $F^*_{v^{(q)}}$ to be the formula computing the X_1-degree-$(r-q)$ homogeneous component of F^*_v.

Thus, $\widehat{F^*_{v^{(q)}}} = \left(\widehat{F^*_v}\right)^{(r-q)}$. Using the definition of F^*_v and Lemma B.19, we obtain $\widehat{F^*_{v^{(q)}}} = (A_F[v, t])^{(r-q)} = A'_F[v^{(q)}, t^{(r)}]$.

Hence every node of the compressed ABP A'_F is represented by a formula obtained by homogenizing an induced formula of F. In particular, these formulas are derivable from the original formula inside Frege by the same homogenization reasoning used earlier.

Lemma B.20. *For every node $v^{(q)}$ of A'_F, the induced formula $F^*_{v^{(q)}}$ has size at most $s^{O(\log^2 s)}$.*

Proof. By [LTW18], $|F^*_v| \le s^{O(\log s)}$. Furthermore, $\mathsf{depth}(F^*_v) = O(\log s)$, since F^*_v is an induced formula of the balanced formula F.

Applying Raz's homogenization to F^*_v produces every homogeneous component with size $|F^*_v|^{O(\mathsf{depth}(F^*_v))}$. Therefore, $|F^*_{v^{(q)}}| \le \left(s^{O(\log s)}\right)^{O(\log s)} = s^{O(\log^2 s)}$. Thus every node of the compressed partially homogeneous ABP can be represented by a quasi-polynomial-size formula derivable inside Frege. □

B.4.3 Local identities for the partially homogenized ABP

Fix r and continue to write A'_F for the compressed ABP $A'_{F,r}$. By construction, every edge from layer q to layer $q+1$ has a label $G^{(q)}_{v,u} = \sum_{j=1}^{n_1} x_{1,j} f^{(q)}_{v,u,j}$, where each $f^{(q)}_{v,u,j} \in \mathbb{F}\langle X_2 \rangle$ is computed by a polynomial-size noncommutative ABP.

Since $\widehat{F^*_{v^{(q)}}} = A'_F[v^{(q)}, t^{(r)}]$, the recursive computation in the compressed ABP gives the following local identity.

Lemma B.21 (Local identity for A'_F)**.** *Let $v^{(q)}$ be a node of A'_F with $0 \le q < r$. Then*

$$F^*_{v^{(q)}} = \sum_{u:(v^{(q)},u^{(q+1)}) \in E(A'_F)} G^{(q)}_{v,u} \cdot F^*_{u^{(q+1)}},$$

where at the final layer we regard $F^*_{t^{(r)}} = 1$.

Proof. Semantically, the identity follows immediately from the recursive computation of the ABP A'_F: $A'_F[v^{(q)}, t^{(r)}] = \sum_{u:(v^{(q)},u^{(q+1)}) \in E(A'_F)} G^{(q)}_{v,u} \cdot A'_F[u^{(q+1)}, t^{(r)}]$. Substituting $A'_F[u^{(q+1)}, t^{(r)}] = \widehat{F^*_{u^{(q+1)}}}$ gives the required algebraic identity.

We additionally note that this identity is obtained syntactically from the local identities of the intermediate ABP $\widetilde{A}_F$. Indeed, starting from the local recursion at $v^{(q)}$, repeatedly substitute through degree-preserving X_2-edges until the next X_1-edge is reached. Since $\widetilde{A}_F$ is acyclic, this procedure terminates after polynomially many such substitutions. The resulting coefficients are exactly the path polynomials $C^{(q)}_{v,u}$ used in the definition of the compressed edge labels $G^{(q)}_{v,u}$. Thus the compression preserves not only the computed polynomial but also gives an explicit derivation of the above local identity. □

Writing the identity layer by layer, suppose $v = u_{q,j}$ is a node in the q-th X_1-degree layer of the compressed ABP and its outgoing edges go to $u_{q+1,\ell}$. Then $F^*_{u_{q,j}} = \sum_\ell G^{(q)}_{j,\ell} F^*_{u_{q+1,\ell}}$, where $G^{(q)}_{j,\ell} = \sum_{a=1}^{n_1} x_{1,a} f^{(q)}_{j,\ell,a}$ and every $f^{(q)}_{j,\ell,a}$ is computed by a polynomial-size noncommutative ABP over X_2.

Corollary B.22. *For every node $v^{(q)}$ of A'_F, Frege derives the Booleanization of the local identity from Lemma B.21 with quasi-polynomial-size overhead.*

Proof. The compression from $\widetilde{A}_F$ to A'_F is obtained by polynomially many applications of associativity, distributivity, and substitution through degree-preserving X_2-edges. After Booleanization over $\mathbf{GF}(2)$, these operations correspond to standard Frege reasoning involving XOR and conjunction.

Every coefficient polynomial $C^{(q)}_{v,u}$ and $f^{(q)}_{v,u,j}$ is represented by a polynomial-size noncommutative ABP over X_2. Using the standard divide-and-conquer conversion, each such ABP is represented by a quasi-polynomial-size noncommutative formula. Whenever two such representa-

tions compute the same polynomial, their equality is a noncommutative identity over X_2, and its Booleanization has a quasi-polynomial-size Frege proof by [LTW18].

Therefore Frege can formalize each step of the compression and derive the Booleanization of $F^*_{v^{(q)}} = \sum_u G^{(q)}_{v,u} F^*_{u^{(q+1)}}$ with quasi-polynomial-size overhead. Since the number of nodes, edges, and compression steps is polynomial in the size of $\widetilde{A}_F$, the total proof size remains quasi-polynomial in the size of the original formula F. □

B.5 Identity Witnesses

Note that after the partial homogenization procedure and refined tracking (Subsection B.3, Subsection B.4) on the partially commutative formula $F \in \mathbb{F}\langle X_1 \sqcup X_2\rangle$, we obtain an ABP A'_F that is homogeneous over X_1 and every coefficient has polynomial size non commutative ABP over X_2. For each nonzero X_1-degree $r \in \{1, \ldots, d\}$, the construction gives such an ABP $A'_{F,r}$; the X_1-degree-zero component is a noncommutative identity over X_2 and is handled directly by [LTW18]. Fix one nonzero-degree partially homogeneous ABP. For simplicity, we drop the X_1-degree index and denote it by A'_F.

Using the refined operator D, for every node at every layer $v = u^{(r)}$ of A'_F, the induced formula F^*_v computes $A'_F[v, t]$. This formula computes a polynomial in $\mathbb{F}\langle X_1 \sqcup X_2\rangle$ that is homogeneous over X_1.

We now apply the identity witness lemma, which is restated below. The proof of this identity follows from our white-box identity testing algorithm and a Factorization Lemma (Lemma 3.5) and is given in Subsection 3.4.

Lemma 1.6. (Witness Theorem for Partially Commutative ABP Identities) *Let F be a field and A be a partially commutative homogeneous (over X_1) ABP in $\mathbb{F}\langle X_1 \sqcup X_2\rangle$ of size s computing the identically zero polynomial. Let d be the depth of A, and let w_i denote the width of the i-th layer. Then there exist $2d-1$ matrices $\{\lambda'_0, \ldots, \lambda'_{d-1}\}$ and $\{T'_1, \ldots, T'_{d-1}\}$ such that*

$$\lambda'_i \in \mathbb{F}^{w_i \times w_i}\langle X_2\rangle, \qquad T'_i \in \mathbb{F}^{w_{i-1} \times w_i}\langle X_1 \sqcup X_2\rangle,$$

and the following hold:

$$\lambda'_{d-i} \cdot \overline{A}_i = \overline{0} \quad \forall\, i = 1, 2, \ldots, d, \tag{1.6}$$

$$\lambda'_{d-i} \cdot \overline{A}_i = T'_{d-i+1} \cdot \lambda'_{d-i+1} \cdot \overline{A}_{i-1} \quad \forall\, i = 2, 3, \ldots, d, \tag{1.7}$$

$$\lambda'_0 := 1, \qquad \lambda'_0 \cdot \overline{A}_d = \overline{A}_d = 0. \tag{1.8}$$

Moreover, every entry of the matrices λ'_i and T'_j is computable by a polynomial-size noncommutative ABP.

B.6 The Frege Simulation

Throughout this section, the field is $\mathbb{F} = \mathbf{GF}(2)$.

Let F be a fully homogeneous, partially commutative formula over $\mathbb{F}\langle X_1 \sqcup X_2\rangle$, and let A_F be the corresponding homogeneous ABP obtained by Nisan's construction. For every $r \in [d]$, let $A'_{F,r}$ be the compressed partially homogeneous ABP obtained from A_F by the construction of Lemma B.17; its sink is denoted by $t^{(r)}$. For a fixed r, we write $v^{(p)}$ for a node in the X_1 degree-p layer of $A'_{F,r}$.

By the previous subsection, every node $v^{(p)}$ of A'_F is associated with a formula $F^*_{v^{(p)}}$ satisfying

$$\widehat{F^*_{v^{(p)}}} = A'_{F,r}[v^{(p)}, t^{(r)}].$$

Moreover, every formula $F^*_{v^{(p)}}$ has size at most $s^{O(\log^2 s)}$ and admits a quasi-polynomial size Frege derivation from the original formula F. The X_1-degree-zero component is treated separately as a noncommutative formula over X_2.

We are finally ready to prove Theorem B.2, which we recall first.

Theorem B.2. *Let $X = X_1 \sqcup X_2$ be partially commutative variables, where $|X_1| = n_1$, $|X_2| = n_2$, and $N = n_1 + n_2$. Let F be a partially commutative formula of size s over $\mathbb{F}$, with variable set X, that computes the zero polynomial. Then there exists a constant $c > 0$ such that the Boolean formula $\neg\widetilde{F}(X)$ has a* Frege *proof of size $s^{(\log s)^c}$.*

Booleanization

Following Definition 3.1, for every algebraic formula P over $\mathbb{F}$ we denote by $\widetilde{P}$ its Booleanization, obtained by replacing $+$ with $\oplus$ and $\cdot$ with $\wedge$.

Booleanization of the homogeneous decomposition

Let s and t denote the source and sink of the homogeneous ABP A_F. After partial homogenization with respect to X_1, the polynomial $A_F[s,t]$ decomposes into its X_1-homogeneous components of degrees $0, 1, \ldots, d$. For every $p \in \{1, \ldots, d\}$, let $A'_{F,p}$ be the compressed ABP from Subsection B.4, with sink $t^{(p)}$, and let $F^*_{s,p}$ be the induced formula associated with its source. By Lemma B.19,

$$\widehat{F^*_{s,p}} = A'_{F,p}[s, t^{(p)}] = \big(A_F[s,t]\big)^{(p)}.$$

For $p = 0$, let $F^*_{s,0}$ denote the X_1-degree-zero component formula; it is a noncommutative formula over X_2. By the correctness of the partial homogenization,

$$A_F[s,t] = \sum_{p=0}^{d} A'_{F,p}[s, t^{(p)}]. \quad \text{Consequently,} \quad F = \sum_{p=0}^{d} F^*_{s,p}.$$

Applying Booleanization yields

$$\widetilde{F} \equiv \bigoplus_{p=0}^{d} \widetilde{F^*_{s,p}}. \tag{B.23}$$

Claim B.24. *Frege admits a polynomial-size proof of Equation B.23. Moreover,*

$$\bigwedge_{p=0}^{d} \neg \widetilde{F^*_{s,p}} \vdash \neg \widetilde{F}.$$

Proof. By the correctness of the partial homogenization and the construction of the formulas $F^*_{s,p}$, we have $F = \sum_{p=0}^{d} F^*_{s,p}$. Booleanizing this identity yields Equation B.23. Since Booleanization over $\mathbf{GF}(2)$ replaces addition with XOR, Frege derives this equivalence with a polynomial overhead from the local homogenization identities.

Assume that $\bigwedge_{p=0}^{d} \neg \widetilde{F^*_{s,p}}$ holds. Then every summand on the right-hand side of Equation B.23 is equivalent to $\bot$. Using the associativity of XOR together with the propositional equivalence $A \oplus B \equiv (A \wedge \neg B) \vee (\neg A \wedge B)$, Frege derives that

$$\bigoplus_{p=0}^{d} \widetilde{F^*_{s,p}} \equiv \bot.$$

Finally, by Equation B.23, Frege concludes $\neg \widetilde{F}$. The proof size is polynomial in the size of Equation B.23. □

Booleanization of the witness identities

Fix $r \in [d]$ and consider the compressed ABP $A'_{F,r}$ of depth r, with source s and sink $t^{(r)}$. Since

$$\widehat{F^*_{s,r}} = A'_{F,r}[s, t^{(r)}],$$

we may apply the witness theorem (Lemma 1.6) to this ABP.. Hence, there exist witness matrices

$$\lambda'_{r,0}, \ldots, \lambda'_{r,r-1}, \qquad T'_{r,1}, \ldots, T'_{r,r-1},$$

whose entries are computed by polynomial-size noncommutative ABPs and which satisfy

$$\lambda'_{r,r-i} \cdot \overline{A}_i^{(r)} = \overline{0}, \qquad 1 \le i \le r, \tag{B.25}$$

$$\lambda'_{r,r-i} \cdot \overline{A}_i^{(r)} = T'_{r,r-i+1} \cdot \lambda'_{r,r-i+1} \cdot \overline{A}_{i-1}^{(r)}, \qquad 2 \le i \le r. \tag{B.26}$$

Here $\overline{A}_i^{(r)}$ is the vector of partial ABPs from the nodes in layer $r - i$ to the sink $t^{(r)}$.

By Lemma B.19, each entry of $\overline{A}_i^{(r)}$ is computed by the corresponding induced formula $F^*_{u^{(r)}_{r-i,j}}$. Booleanizing the witness identities therefore gives the following propositional identities.

Vanishing identities. For every layer i and every row index a,

$$\bigoplus_j \left(\widetilde{\lambda'_{r,r-i}[a,j]} \wedge \widetilde{F^*_{u^{(r)}_{r-i,j}}} \right) \equiv \bot. \tag{B.27}$$

Transition identities. For every $2 \le i \le r$ and every row index a,

$$\bigoplus_j \left(\widetilde{\lambda'_{r,r-i}[a,j]} \wedge \widetilde{F^*_{u^{(r)}_{r-i,j}}} \right) \equiv \bigoplus_j \left(\widetilde{T'_{r,r-i+1}[a,j]} \wedge \bigoplus_\ell \left(\widetilde{\lambda'_{r,r-i+1}[j,\ell]} \wedge \widetilde{F^*_{u^{(r)}_{r-i+1,\ell}}} \right) \right). \tag{B.28}$$

Key Frege reflection step.

Claim B.29. *Each Booleanized identity in Equation B.27 and Equation B.28 has a quasi-polynomial size Frege proof.*

Proof. We prove that each Booleanized identity in Equation B.27 and Equation B.28 admits a Frege proof of size quasi-polynomial in $|A'_F|$.

Fix $r \in [d]$ and write $A'_{F,r}$ for the compressed ABP of depth r. Now consider each ABP $A'_{F,r}[s,t^{(r)}]$ where t was the sink node in A_F, and $A'_{F,r}[s,t^{(r)}]$ computes the X_1-degree r homogeneous component of A_F. For each such r, we have a set of witness identities, denoted as $\left\{\lambda'_{r,i}, T'_{r,i}\right\}$. Moreover, we have the length of $A'_{F,r}$ is r, which is the total degree and assume the width at layer i, $w_{r,i}$.

Vanishing identities Fix $r \in [d]$. We consider the identity $\lambda'_{r,r-i} \cdot \overline{A}_i^{(r)} = 0$. We first treat the base layer $i = 1$.

Base case ($i = 1$). Each entry of $\overline{A}_1^{(r)}$ is a homogeneous linear form in the X_1-variables with coefficients in $\mathbb{F}\langle X_2 \rangle$. Thus, for every $t \in [w_{r,r-1}]$,

$$\overline{A}_1^{(r)}[t] = \sum_{j=1}^{n_1} x_{1,j} \cdot f_{t,j}, \qquad f_{t,j} \in \mathbb{F}\langle X_2 \rangle.$$

Therefore, the identity for every $p \in w_{r,r-1}$

$$\sum_t \lambda'_{r,r-1}[p,t] \cdot \overline{A}_1^{(r)}[t] = 0$$

is equivalent to the collection of identities

$$\forall j \in [n_1], \qquad \sum_t \lambda'_{r,r-1}[p,t] \cdot f_{t,j} = 0.$$

Each polynomial $\sum_t \lambda'_{r,r-1}[p,t] \cdot f_{t,j}$ is computed by a noncommutative ABP over X_2 of size $\mathrm{poly}(|A'_F|)$. Hence, by [LTW18], each Boolean tautology

$$\neg \bigoplus_t \left(\widetilde{\lambda'_{r,r-1}[p,t]} \wedge \widetilde{f}_{t,j} \right)$$

has a quasi-polynomial size Frege proof. Combining these for all j and substituting the local induced-formula identities, Frege derives the Booleanized identity Equation B.27 for $i = 1$.

General i. Assume that the Booleanized transition identities (Equation B.28) have already been proved for every layer. We show that the Booleanized vanishing identities follow.

The witness identities satisfy

$$\lambda'_{r,r-i} \cdot \overline{A}_i^{(r)} = T'_{r,r-i+1} \cdot \lambda'_{r,r-i+1} \cdot \overline{A}_{i-1}^{(r)} \qquad \forall\, 2 \le i \le r.$$

By the base case proved above,

$$\lambda'_{r,r-1} \cdot \overline{A}_1^{(r)} = 0.$$

Suppose inductively that

$$\lambda'_{r,r-i+1} \cdot \overline{A}_{i-1}^{(r)} = 0.$$

Substituting this identity into the transition identity yields

$$\lambda'_{r,r-i} \cdot \overline{A}_i^{(r)} = T'_{r,r-i+1} \cdot 0 = 0.$$

Applying Booleanization, Frege derives the corresponding vanishing identity at layer i. Therefore, once the transition identities are proved, the Booleanized vanishing identities for every layer follow by induction on i.

Transition identities Fix $i \in \{2, \dots, r\}$. We consider the identity

$$\lambda'_{r,r-i} \cdot \overline{A}_i^{(r)} = T'_{r,r-i+1} \cdot \lambda'_{r,r-i+1} \cdot \overline{A}_{i-1}^{(r)}.$$

Since this is a vector identity, it suffices to prove it entry-wise. Fix a row index p. Then the above identity is equivalent to

$$\sum_{t \in [w_{r,r-i}]} \lambda'_{r,r-i}[p,t] \cdot \overline{A}_i^{(r)}[t] = \sum_{t \in [w_{r,r-i+1}]} T'_{r,r-i+1}[p,t] \cdot \left(\sum_{q \in [w_{r,r-i+1}]} \lambda'_{r,r-i+1}[t,q] \cdot \overline{A}_{i-1}^{(r)}[q] \right).$$

We now reduce this identity to noncommutative ABP identities over X_2. Since the ABP is homogeneous over X_1, the layer transition satisfies

$$\overline{A}_i^{(r)} = \left(\sum_{j=1}^{n_1} x_{1,j} M_j^{(r,i-1)} \right) \cdot \overline{A}_{i-1}^{(r)},$$

where each $M_j^{(r,i-1)}$ is a matrix with entries in $\mathbb{F}\langle X_2 \rangle$ computed by polynomial-size ABPs.

Similarly, by construction of the witness matrices (see Lemma 1.6), the transition matrix admits a linear expansion

$$T'_{r,r-i+1} = \sum_{j=1}^{n_1} x_{1,j} C_j^{(i-1)},$$

where each $C_j^{(i-1)}$ has entries in $\mathbb{F}\langle X_2 \rangle$ and is computable by polynomial-size ABPs.

Substituting these expressions, the identity becomes

$$\lambda'_{r,r-i} \cdot \left(\sum_{j=1}^{n_1} x_{1,j} M_j^{(r,i-1)} \right) \cdot \overline{A}_{i-1}^{(r)} = \left(\sum_{j=1}^{n_1} x_{1,j} C_j^{(i-1)} \right) \cdot \lambda'_{r,r-i+1} \cdot \overline{A}_{i-1}^{(r)}.$$

Since the variables $x_{1,1}, \ldots, x_{1,n_1}$ appear as the leftmost noncommuting variables, it therefore suffices to establish the following coefficient-wise identities for each $x_{1,j} \quad \forall j \in [n_1]$. That is,

$$\left(\lambda'_{r,r-i} M_j^{(r,i-1)}\right) \cdot \overline{A}_{i-1}^{(r)} = \left(C_j^{(i-1)} \lambda'_{r,r-i+1}\right) \cdot \overline{A}_{i-1}^{(r)}.$$

By Lemma 1.6, the matrices $C_j^{(i-1)}$ are constructed so that the following matrix identities hold over $\mathbb{F}\langle X_2 \rangle$:

$$\lambda'_{r,r-i} M_j^{(r,i-1)} = C_j^{(i-1)} \lambda'_{r,r-i+1} \qquad \forall j \in [n_1].$$

Expanding these identities entry-wise, for each $j \in [n_1]$ and each pair (a, b), we obtain noncommutative polynomial identities

$$\left(\lambda'_{r,r-i} M_j^{(r,i-1)}\right)[a,b] = \left(C_j^{(i-1)} \lambda'_{r,r-i+1}\right)[a,b].$$

ach side is computed by a polynomial-size noncommutative ABP over X_2. Hence, by [LTW18], the Booleanization of each such identity admits a quasi-polynomial size Frege proof. Since the total number of such identities is at most $n_1 \cdot w^2 \leq Nw^2$ (where $w = \max_i w_{r,i}$), Frege can derive all these Booleanized identities with a quasi-polynomial overhead. Using standard propositional reasoning expressing matrix multiplication entry-wise over $\mathbb{F}$, Frege combines these identities to derive the Booleanization of the matrix identities

$$\lambda'_{r,r-i} M_j^{(r,i-1)} = C_j^{(i-1)} \lambda'_{r,r-i+1} \quad \forall j \in [n_1].$$

Substituting these identities back into the coefficient-wise equalities, Frege derives the Booleanization of

$$\left(\lambda'_{r,r-i} M_j^{(r,i-1)}\right) \cdot \overline{A}_{i-1}^{(r)} = \left(C_j^{(i-1)} \lambda'_{r,r-i+1}\right) \cdot \overline{A}_{i-1}^{(r)}.$$

Finally, recombining over all $j \in [n_1]$ using the linear expansion in X_1, Frege derives the Booleanized transition identity Equation B.28.. □

Corollary B.30 (Frege proofs for partially homogeneous ABP identities)**.** *For every $r \in \{0, \ldots, d\}$, if the X_1-homogeneous component of degree r computes the zero polynomial, then the Booleanized formula $\neg\widetilde{F^*_{s,r}}$ admits a quasi-polynomial-size Frege proof.*

Proof. If $r = 0$, then $F^*_{s,0}$ is a noncommutative formula over X_2, so the claim follows directly from [LTW18]. Assume $r \geq 1$. By the Witness Theorem (Lemma 1.6), the compressed ABP $A'_{F,r}[s, t^{(r)}]$ admits witness matrices $\{\lambda'_{r,i}\}_{i=0}^{r-1}$ and $\{T'_{r,i}\}_{i=1}^{r-1}$ that satisfy the vanishing and transition identities. By Claim B.29, the Booleanizations of these identities have quasi-polynomial-size Frege proofs. Applying the layer-by-layer propagation argument below, Frege derives $\neg\widetilde{F^*_{s,r}}$. □

We are finally ready to combine the above ideas to write a Frege derivation as required in Theorem B.2.

Theorem B.2. *Let $X = X_1 \sqcup X_2$ be partially commutative variables, where $|X_1| = n_1$, $|X_2| = n_2$, and $N = n_1 + n_2$. Let F be a partially commutative formula of size s over $\mathbb{F}$, with variable set X, that computes the zero polynomial. Then there exists a constant $c > 0$ such that the Boolean formula $\neg\widetilde{F}(X)$ has a* Frege *proof of size $s^{(\log s)^c}$.*

Proof. As in [LTW18], we first balance the original size-s formula to logarithmic depth and apply Lemma B.3. Since F computes the zero polynomial, each fully homogeneous component also computes zero, and Equation B.4 shows that it suffices to prove the desired Frege statement for each such component. Thus, we may assume, as throughout this subsection, that F is fully homogeneous.

By Claim B.24, it suffices to derive $\neg\widetilde{F^*_{s,r}}$ for every X_1-homogeneous component. The degree-zero component $r = 0$ is a noncommutative identity over X_2 and is handled directly by [LTW18]. Fix $r \in \{1, \ldots, d\}$. Throughout the remainder of the layer-by-layer argument, we consider the compressed partially homogeneous ABP $A'_{F,r}[s, t^{(r)}]$. For simplicity, we suppress the superscript (r) from the witness matrices.

For every layer-distance $i \in [r]$ and every row index a of $\lambda'_{r,r-i}$, define

$$\Phi_i(a) = \bigoplus_t \left(\widetilde{\lambda'_{r,r-i}[a,t]} \wedge \widetilde{F^*_{u^{(r)}_{r-i,t}}} \right).$$

By Lemma B.19, this is precisely the Booleanization of the a-th coordinate of $\lambda'_{r,r-i}\overline{A}_i^{(r)}$.

Base case. By Equation B.27, Frege derives $\neg\Phi_1(a)$ for every row index a.

Induction step. Assume Frege has derived $\neg\Phi_{i-1}(t)$ for every relevant row index t. By Equation B.28, Frege derives

$$\Phi_i(a) \leftrightarrow \bigoplus_t \left(\widetilde{T'_{r,r-i+1}[a,t]} \wedge \Phi_{i-1}(t) \right).$$

Every conjunction on the right is false, so Frege derives $\neg\Phi_i(a)$. Thus, the assertion propagates through all r layers. At $i = r$, the backward vector is the single source-to-sink polynomial and $\lambda'_{r,0} = 1$. Hence,

$$\Phi_r(1) = \widetilde{F^*_{s,r}},$$

and Frege derives $\neg\widetilde{F^*_{s,r}}$.

Together with the degree-zero case handled directly by [LTW18], Frege therefore derives

$$\neg\widetilde{F^*_{s,r}} \qquad \forall r \in \{0,\dots,d\}.$$

Finally, Claim B.24 yields $\neg\widetilde{F}$.

Proof size. Let s denote the size of the original partially commutative formula F. The construction first homogenizes F, then converts the resulting formula into a partially commutative ABP, and finally partially homogenizes the ABP with respect to X_1. Let

$$N' := \max_{1\le r\le d} |A'_{F,r}|$$

denote the maximum size of a compressed partially homogeneous ABP (and enlarge N' if necessary to dominate the size of the degree-zero component formula). By the previous size bounds, we have $N' \le s^{O(\log s)}$. In particular, $\log N' = O(\log^2 s)$. The depth and width of every $A'_{F,r}$ are at most N'.

We first bound the size of the formulas occurring in the Frege proof. By the construction of the induced formulas, every formula F^*_v has size at most $s^{O(\log^2 s)}$. Equivalently, in terms of N', this is bounded by $N'^{O(\log N')}$. The entries of the witness matrices $\left\{\lambda'_{r,i} \,:\, r \in [1,\dots,d], i \in [0,\dots,r-1]\right\}$ and $\left\{T'_{r,i} \,:\, r \in [1,\dots,d], i \in [1,\dots,r-1]\right\}$ are computed by polynomial-sized noncommutative ABPs. Using the standard ABP-to-formula conversion, each such entry can be represented by a noncommutative formula of size $N'^{O(\log N')}$. After Booleanization, these become Boolean formulas of the same quasi-polynomial size.

We next consider the Booleanized witness identities. For every fixed r, the proof of Claim B.29 reduces these identities to a polynomial number of noncommutative ABP identities over X_2. The number of such identities is polynomial in N'. Each of these identities is computed by noncommutative formulas of quasi-polynomial size in N'. Hence, by the Frege simulation of noncommutative

identity testing in [LTW18], each identity has a Frege proof of size $N'^{O(\log^{O(1)} N')}$. Therefore, since only polynomially many such identities are used, the total size contributed by the NC identity proofs is still $N'^{O(\log^{O(1)} N')}$.

It remains to account for the layer-by-layer propagation. For each fixed X_1-degree r, there are at most d layers and at most N rows in each layer. Thus, the propagation uses at most $O(dN') \leq O(N'^2)$ steps. Each step uses one Booleanized witness identity and a polynomial number of propositional substitution and simplification steps. Hence, the additional number of Frege lines needed for the propagation is polynomial in N'.

There are at most $d + 1 \leq N' + 1$ values of r. Thus, combining the proofs for all X_1-homogeneous components introduces only another polynomial factor in N'. Consequently, the complete Frege derivation has $N'^{O(\log^{O(1)} N')}$ lines, and every line has a size of at most $N'^{O(\log N')}$. Therefore, the total Frege proof size is $N'^{O(\log^{O(1)} N')}$. Since $N \leq s^{O(\log s)}$, we have $\log N' = O(\log^2 s)$. Hence, the Frege proof of $\neg\widetilde{F}$ has size $s^{(\log s)^{O(1)}}$. Thus, the Booleanization of the partially commutative identity $F = 0$ admits a quasi-polynomial-size Frege proof in the size s of the original formula. □

C Proof of Lemma 3.7

We first state some definitions that we need for the proof.

Definition C.1 (Free-Skew Field). *[Coh95] For a field $\mathbb{F}$ and a set of noncommuting variables X, the free skew field $\mathbb{F}\langle\langle X\rangle\rangle$ is the smallest field containing the polynomial ring $\mathbb{F}\langle X\rangle$.* ◊

We now recall the definition of shrunk subspaces used in the context of the noncommutative rank computation [FR04, IQS18].

Definition C.2. *[Shrunk Subspace] Let $L = \sum_{i=1}^{n} x_i \cdot A_i \in \mathbb{F}^{k\times m}\langle X\rangle$ be a noncommutative linear matrix, where $A_i \in \mathbb{F}^{k\times m}$, and suppose that $\operatorname{ncrank}(L) = r < \min\{k, m\}$. A* shrunk subspace *for L is a subspace $T \subseteq \mathbb{F}^m$ for which there exists a subspace $W \subseteq \mathbb{F}^k$ such that $A_i \cdot T \subseteq W$ for every $i \in [n]$ and $\dim(T) > \dim(W)$. More generally, T is called a c-shrunk subspace if one can choose W with $\dim(T) - \dim(W) \geq c$.* ◊

Next, we recall the notion of r-decomposability of linear matrices from [FR04].

Definition C.3. *[r-Decomposable] A linear matrix $L = \sum_{i=1}^{n} x_i A_i \in \mathbb{F}^{k\times m}\langle X\rangle$ is called* r-decomposable *if there exist invertible matrices $U \in \mathbb{F}^{k\times k}$ and $V \in \mathbb{F}^{m\times m}$ such that*

$$UL[X]V = \left[\begin{array}{c|c} M[X] & 0 \\ \hline B[X] & C[X] \end{array}\right], \tag{C.4}$$

where the block $B[X]$ has dimension $p \times q$ with $p + q = r$, $\operatorname{ncrank}(L) = r$, and M, B, C being noncommutative linear matrices. ◊

Given a linear matrix M, the following equivalent characterizations are known from the work of [Coh95, FR04, IQS17, GGdOW20].

Theorem C.5. (Noncommutative rank of a linear matrix). *Let $L = \sum_{i=1}^{n} A_i x_i \in \mathbb{F}^{\ell \times m} \langle X \rangle$ be a linear matrix and $\ell \geq m$. Then the following are equivalent:*

- *The noncommutative rank of L is r.*
- *There exist linear matrices $L_1 \in \mathbb{F}\langle X \rangle^{\ell \times r}$ and $L_2 \in \mathbb{F}\langle X \rangle^{r \times m}$ such that $L = L_1 \cdot L_2$.*
- *L has an $(m-r)$-*shrunk subspace*, i.e. there exist a subspace $T \subseteq \mathbb{F}^m$ and a subspace $W \subseteq \mathbb{F}^\ell$ such that $A_i T \subseteq W$ for every $i \in [n]$ and $\dim(T) - \dim(W) \geq m - r$.*
- *L is r-decomposable, i.e. there exist invertible matrices $U \in \mathbb{F}^{\ell \times \ell}$ and $V \in \mathbb{F}^{m \times m}$ such that*

$$ULV = \left[\begin{array}{c|c} L' & 0 \\ \hline B & C \end{array}\right], \tag{C.6}$$

where the zero block is of size $i \times j$ with $i + j = \ell + m - r$.

The existence of a shrunk subspace is equivalent to the fact the noncommutative rank is not full.

The classic result of Fortin and Reutenauer [FR04] gives a necessary and sufficient condition for the decomposability of linear matrices. We state their result formally. Then we use their result to obtain an *explicit* matrix factorization of linear matrices.

Theorem C.7 (Fortin- Reutenauer [FR04])**.** *Let $L = \sum_{i=1}^{n} x_i \cdot A_i \in \mathbb{F}^{\ell \times m} \langle X \rangle$ be a noncommutative linear matrix. Then L is r-decomposable if and only if* $\mathrm{ncrank}(L) = r$

A linear matrix L is r-decomposable if and only if it has a shrunk subspace using Theorem C.5.

We divide the proof of Lemma 3.7 into two parts. Given a linear matrix L, we first use [IQS18] to construct a nontrivial shrunk subspace. We then exploit this subspace to obtain a zero block following [FR04], which in turn yields a factorization of the input matrix.

In the following lemma, we first show the construction of a nontrivial subspace.

Lemma C.8. *Let $L = \sum_{i=1}^{n} A_i x_i \in \mathbb{F}\langle X \rangle^{\ell \times m}$ be a noncommutative linear matrix with* $\mathrm{ncrank}(L) = r < \min\{\ell, m\}$*. Then we can find a basis of an $(m-r)$-shrunk subspace $T \subseteq \mathbb{F}^m$.*

Proof. We reduce the rectangular case to the square case. If $\ell < m$, pad L with $m - \ell$ zero rows: for every $i \in [n]$, let

$$\widehat{A}_i = \begin{bmatrix} A_i \\ 0_{(m-\ell)\times m} \end{bmatrix} \in \mathbb{F}^{m \times m}, \qquad \widehat{L} = \sum_{i=1}^{n} \widehat{A}_i x_i.$$

Then $\mathrm{ncrank}(\widehat{L}) = r$. Applying the square noncommutative-rank algorithm of [IQS18] gives an $(m-r)$-shrunk subspace $T \subseteq \mathbb{F}^m$. Since every image $\widehat{A}_i T$ lies in the first ℓ coordinates, the same

subspace T is an $(m-r)$-shrunk subspace for the original matrix L, after identifying the corresponding image space with a subspace of $\mathbb{F}^\ell$. This settles the case $\ell < m$.

Assume henceforth that $\ell \geq m$. For every $i \in [n]$, define $\widetilde{A}_i = \left[\begin{array}{c|c} A_i & 0_{\ell\times(\ell-m)} \end{array}\right] \in \mathbb{F}^{\ell\times\ell}$, and let $\widetilde{L} = \sum_{i=1}^n \widetilde{A}_i x_i = \left[\begin{array}{c|c} L & 0 \end{array}\right]$. Clearly, adding zero columns does not change the noncommutative rank, and hence $\operatorname{ncrank}(\widetilde{L}) = \operatorname{ncrank}(L) = r$.

We now apply the noncommutative rank algorithm of [IQS18] to the square linear matrix $\widetilde{L}$. For sufficiently large fields this follows from [IQS18, Theorem 1.5], while for small finite fields, and in particular for $\mathbb{F} = \mathbf{GF}(2)$, we use [IQS18, Corollary 1.7]. Thus, in deterministic polynomial time we obtain an $(\ell-r)$-shrunk subspace $\widetilde{T} \subseteq \mathbb{F}^\ell$, together with a basis of $\widetilde{T}$ over the ground field $\mathbb{F}$.

Since $\widetilde{T}$ is an $(\ell-r)$-shrunk subspace, by definition there exists a subspace $\widetilde{W}_0 \subseteq \mathbb{F}^\ell$ such that $\widetilde{A}_i\widetilde{T} \subseteq \widetilde{W}_0$ for every $i \in [n]$ and $\dim(\widetilde{T}) - \dim(\widetilde{W}_0) \geq \ell - r$. We now replace this arbitrary witnessing subspace by the canonical image space $\widetilde{W} := \sum_{i=1}^n \widetilde{A}_i\widetilde{T}$. Clearly, $\widetilde{W} \subseteq \widetilde{W}_0$, and hence $\dim(\widetilde{T}) - \dim(\widetilde{W}) \geq \dim(\widetilde{T}) - \dim(\widetilde{W}_0) \geq \ell - r$. On the other hand, since $\operatorname{ncrank}(\widetilde{L}) = r$, the maximum possible shrinkage is $\ell - r$. Therefore, $\dim(\widetilde{T}) - \dim(\widetilde{W}) = \ell - r$. Let $Z := \{0\}^m \oplus \mathbb{F}^{\ell-m} = \operatorname{span}_{\mathbb{F}}\{e_{m+1}, \ldots, e_\ell\} \subseteq \mathbb{F}^\ell$. The subspace Z corresponds exactly to the $\ell - m$ zero columns introduced in the padding. Hence, $\widetilde{A}_i Z = 0$ for every $i \in [n]$.

We claim that $Z \subseteq \widetilde{T}$. Suppose otherwise. Then $\widetilde{T}' := \widetilde{T} + Z$ satisfies $\dim(\widetilde{T}') > \dim(\widetilde{T})$. On the other hand, since every $\widetilde{A}_i$ annihilates Z, $\sum_{i=1}^n \widetilde{A}_i\widetilde{T}' = \sum_{i=1}^n \widetilde{A}_i\widetilde{T} = \widetilde{W}$. Consequently, $\dim(\widetilde{T}') - \dim(\widetilde{W}) > \ell - r$, which contradicts the fact that the maximum shrinkage of $\widetilde{L}$ is $\ell - r$. Therefore, $Z \subseteq \widetilde{T}$.

Consider now the canonical projection $\pi : \mathbb{F}^\ell \to \mathbb{F}^m$ onto the first m coordinates. Since $\ker(\pi) = Z$, define $T := \pi(\widetilde{T}) \subseteq \mathbb{F}^m$. As $Z \subseteq \widetilde{T}$, we have $T \cong \widetilde{T}/Z$, and therefore $\dim(T) = \dim(\widetilde{T}) - \dim(Z) = \dim(\widetilde{T}) - (\ell - m)$.

Let $W := \sum_{i=1}^n A_i T \subseteq \mathbb{F}^\ell$. We claim that $W = \widetilde{W}$. In fact, we show the stronger statement $A_i T = \widetilde{A}_i\widetilde{T}$ for every $i \in [n]$.

Recall that every vector $\widetilde{t} \in \mathbb{F}^\ell$ can be written as $\widetilde{t} = (t, z)$ with $t \in \mathbb{F}^m$ and $z \in \mathbb{F}^{\ell-m}$, and that $\pi(\widetilde{t}) = t$. Since $\widetilde{A}_i = \left[\begin{array}{c|c} A_i & 0 \end{array}\right]$, we have $\widetilde{A}_i\widetilde{t} = A_i\pi(\widetilde{t})$ for every $\widetilde{t} \in \mathbb{F}^\ell$.

Therefore, if $y \in \widetilde{A}_i\widetilde{T}$, then $y = \widetilde{A}_i\widetilde{t}$ for some $\widetilde{t} \in \widetilde{T}$. Setting $t = \pi(\widetilde{t})$, we have $t \in T$ and hence $y = A_i t \in A_i T$. This proves $\widetilde{A}_i\widetilde{T} \subseteq A_i T$.

Conversely, if $y \in A_i T$, then $y = A_i t$ for some $t \in T$. Since $T = \pi(\widetilde{T})$, there exists a lift $\widetilde{t} \in \widetilde{T}$ such that $\pi(\widetilde{t}) = t$. Hence $y = A_i t = A_i\pi(\widetilde{t}) = \widetilde{A}_i\widetilde{t} \in \widetilde{A}_i\widetilde{T}$. Thus $A_i T = \widetilde{A}_i\widetilde{T}$ for every $i \in [n]$, and consequently $W = \sum_{i=1}^n A_i T = \sum_{i=1}^n \widetilde{A}_i\widetilde{T} = \widetilde{W}$. It follows that $\dim(T) - \dim(W) = \dim(\widetilde{T}) - (\ell - m) - \dim(\widetilde{W}) = (\ell - r) - (\ell - m) = m - r$. Thus, $T \subseteq \mathbb{F}^m$ is an $(m-r)$-shrunk subspace for L.

It remains to show that a basis of T can be computed efficiently over the ground field. An explicit basis of Z is $\{e_{m+1}, \ldots, e_\ell\}$. Since $Z \subseteq \widetilde{T}$ and a basis of $\widetilde{T}$ over $\mathbb{F}$ is already available,

Gaussian elimination over $\mathbb{F}$ allows us to extend this basis of Z to a basis $\{e_{m+1}, \dots, e_\ell, v_1, \dots, v_h\}$ of $\widetilde{T}$. Then $\{\pi(v_1), \dots, \pi(v_h)\}$ is a basis of T.

Indeed, these vectors span $\pi(\widetilde{T}) = T$. Moreover, if $\sum_{j=1}^{h} \alpha_j \pi(v_j) = 0$, then $\sum_{j=1}^{h} \alpha_j v_j \in \ker(\pi) = Z$. Since the vectors $v_1, \dots, v_h$ form a complement of Z inside $\widetilde{T}$, this implies $\alpha_1 = \cdots = \alpha_h = 0$. Thus, the projected vectors are linearly independent.

All the above operations, after obtaining $\widetilde{T}$, consist only of ordinary linear algebra over the ground field $\mathbb{F}$ and can be performed in deterministic polynomial time. In particular, the construction works over $\mathbb{F} = \mathbf{GF}(2)$. □

The second part of the proof uses the shrunk subspace obtained to compute the factors. We do this in two parts. First, we compute a nontrivial decomposition, which then implies the factors.

Lemma C.9. *Let $L = \sum_{i=1}^{n} A_i x_i \in \mathbb{F}\langle X \rangle^{\ell \times m}$ be a noncommutative linear matrix with $\operatorname{ncrank}(L) = r < \min\{\ell, m\}$. Let $T \subseteq \mathbb{F}^m$ be a $(m-r)$-shrunk subspace corresponding to L. Then, given a basis of T, one can construct invertible matrices $U \in \mathbb{F}^{\ell \times \ell}$ and $V \in \mathbb{F}^{m \times m}$ such that*

$$ULV = \left[\begin{array}{c|c} M & 0 \\ \hline B & C \end{array}\right], \tag{C.10}$$

where the zero block is of size $a \times b$ such that $a + b = \ell + m - r$.

Proof. We show the existence of invertible matrices $U \in \mathbb{F}^{\ell \times \ell}$ and $V \in \mathbb{F}^{m \times m}$ such that for every $i \in [n]$,

$$UA_iV = \left[\begin{array}{c|c} M_i & 0 \\ \hline B_i & C_i \end{array}\right]. \tag{C.11}$$

Let $\dim(T) = r'$ and let $B_T := \{u_1, \dots, u_{r'}\} \subset \mathbb{F}^m$ be a spanning set of T. We extend B_T to a basis of $\mathbb{F}^m$ by adding $\{u_{r'+1}, \dots, u_m\}$ to B_T and let $T' := \operatorname{span}_{\mathbb{F}}\{u_{r'+1}, \dots, u_m\}$. Thus $T \oplus T' = \mathbb{F}^m$. We define the matrix $V := [u_1, \dots, u_{r'}, u_{r'+1}, \dots, u_m]$ to be the matrix whose columns are the vectors $\{u_1, \dots, u_m\}$. Clearly, V is invertible over $\mathbb{F}$.

Let $W := \sum_{i=1}^{n} A_i T = \operatorname{span}_{\mathbb{F}}\{\bigcup_{i=1}^{n} A_i \cdot T\}$, and write $s = \dim(W)$.

Let $B_W := \{w_1, \dots, w_s\} \subset \mathbb{F}^\ell$ be a basis of W. We extend this basis to a basis of $\mathbb{F}^\ell$ by adding $\{w_{s+1}, \dots, w_\ell\}$ to B_W. Let U be the invertible matrix in $\mathbb{F}^{\ell \times \ell}$ such that $U := [w_1, \dots, w_\ell]^{-1}$. Consider the matrix $A_i \cdot V$. For every $j \leq r'$, the j-th column of this matrix is $A_i \cdot u_j \in W$ by the definition of the shrunk subspace. Hence, for every $j \leq r'$, there exist $\alpha_1, \dots, \alpha_s \in \mathbb{F}$ such that

$A_i u_j = \sum_{i=1}^{s} \alpha_i \cdot w_i$. Formally,

$$U^{-1} \cdot \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_s \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \left[\begin{array}{c|c|c|c|c} w_1 & w_2 & w_3 & \dots & w_\ell \end{array}\right] \cdot \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_s \\ 0 \\ \vdots \\ 0 \end{pmatrix} = A_i u_j \implies U \cdot (A_i \cdot u_j) = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_s \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{C.12}$$

Thus, the last $\ell - s$ positions in $U \cdot (A_i u_j)$ are 0 for every $j \leq r'$.

Therefore, for every $i \in [n]$ we obtain

$$UA_iV = \left[\begin{array}{c|c} * & * \\ \hline 0 & * \end{array}\right], \tag{C.13}$$

where the zero block is of size $(\ell - s) \times r'$. Since T is an $(m-r)$-shrunk subspace, $r' - s \geq m - r$. On the other hand, because $\operatorname{ncrank}(L) = r$, no shrunk subspace can have deficiency greater than $m - r$. Hence $r' - s = m - r$, and therefore $\ell - s + r' = \ell + m - r$.

By permuting the row blocks so that the last $\ell - s$ rows come first, and permuting the column blocks so that the first r' columns come last, we move this zero block to the upper-right corner. Composing these permutations with U and V, we obtain invertible matrices $U' \in \mathbb{F}^{\ell \times \ell}$ and $V' \in \mathbb{F}^{m \times m}$ such that

$$U'A_iV' = \left[\begin{array}{c|c} M_i & 0 \\ \hline B_i & C_i \end{array}\right] \quad \text{for every } i \in [n].$$

Consequently,

$$U'LV' = U' \left(\sum_i x_i A_i \right) V' = \left[\begin{array}{c|c} M & 0 \\ \hline B & C \end{array}\right]. \qquad \square$$

We now show that given the r-decomposability form Equation C.10, the linear factors are easy to compute.

Corollary C.14. *Let $L = \sum_{i=1}^{n} x_i \cdot A_i \in \mathbb{F}\langle X \rangle^{\ell \times m}$ be a noncommutative linear matrix with $\operatorname{ncrank}(L) = r$. Then L admits a nontrivial factorization*

$$L = L_1 \cdot L_2, \qquad L_1 \in \mathbb{F}\langle X \rangle^{\ell \times r}, \quad L_2 \in \mathbb{F}\langle X \rangle^{r \times m},$$

where the factorization follows from the r-decomposability of L.

Proof. Since L is r-decomposable, we have invertible matrices $U \in \mathbb{F}^{\ell\times\ell}$ and $V \in \mathbb{F}^{m\times m}$ such that

$$U \cdot L \cdot V = \left[\begin{array}{c|c} M & 0 \\ \hline B & C \end{array}\right],$$

where the zero block is of size $i \times j$ such that $i + j = \ell + m - r$. Therefore, we can write:

$$L = U^{-1}\left[\begin{array}{c|c} M & 0 \\ \hline B & I_{\ell-i} \end{array}\right] \cdot \left[\begin{array}{c|c} I_{m-j} & 0 \\ \hline 0 & C \end{array}\right] \cdot V^{-1}$$

$$\text{where} \quad L_1 = U^{-1}\left[\begin{array}{c|c} M & 0 \\ \hline B & I_{\ell-i} \end{array}\right] \quad \text{and} \quad L_2 = \left[\begin{array}{c|c} I_{m-j} & 0 \\ \hline 0 & C \end{array}\right] \cdot V^{-1}. \qquad \square$$

Proof of Lemma 3.7. Write the affine-linear matrix as

$$L = A_0 + \sum_{i=1}^{n} A_i x_i.$$

Introduce one additional noncommuting variable x_0 and consider the homogeneous linear matrix

$$L^{\mathrm{h}} = A_0 x_0 + \sum_{i=1}^{n} A_i x_i.$$

Its noncommutative rank is the same as that of L. Indeed, in the free skew field, right multiplication by x_0^{-1} gives

$$L^{\mathrm{h}} x_0^{-1} = A_0 + \sum_{i=1}^{n} A_i (x_i x_0^{-1}),$$

and the rational change of variables $y_i = x_i x_0^{-1}$ is invertible (with inverse $x_i = y_i x_0$). Thus $\operatorname{ncrank}(L^{\mathrm{h}}) = \operatorname{ncrank}(L) = r$.

If $r = 0$, then $L = 0$ and the asserted factorization is trivial. Assume $r > 0$. By Lemma C.8, applied to L^{h}, we can compute a basis of an $(m-r)$-shrunk subspace $T \subseteq \mathbb{F}^m$. Applying Lemma C.9 to this basis gives an r-decomposition of L^{h}. Finally, Corollary C.14 gives linear matrices

$$\widetilde{L}_1 \in \mathbb{F}\langle x_0, X\rangle^{\ell\times r}, \qquad \widetilde{L}_2 \in \mathbb{F}\langle x_0, X\rangle^{r\times m}$$

with $L^{\mathrm{h}} = \widetilde{L}_1 \widetilde{L}_2$. Setting $x_0 = 1$ yields

$$L = L_1 L_2, \qquad L_1 := \widetilde{L}_1|_{x_0=1}, \quad L_2 := \widetilde{L}_2|_{x_0=1},$$

where $L_1 \in \mathbb{F}\langle X\rangle^{\ell\times r}$ and $L_2 \in \mathbb{F}\langle X\rangle^{r\times m}$ are affine-linear matrices. All steps use deterministic polynomial-time linear algebra together with the shrunk-subspace computation, so the construction runs in deterministic $\mathrm{poly}(\ell, m, n)$ time in the arithmetic model over the ground field. $\square$

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