Electronic Colloquium on Computational Complexity, Revision 1 of Report No. 164 (2018)

On the symmetries of and equivalence test for design polynomials

Nikhil Gupta
Indian Institute of Science
nikhilg@iisc.ac.in

Chandan Saha
Indian Institute of Science
chandan@iisc.ac.in

June 24, 2019

Abstract

In a Nisan-Wigderson design polynomial (in short, a design polynomial), every pair of monomials share a few common variables. A useful example of such a polynomial, introduced in [KSS14], is the following:

$$\mathsf{NW}_{d,k}(\mathbf{x}) = \sum_{h \in \mathbb{F}_d[z],\ \deg(h) \le k} \ \prod_{i=0}^{d-1} x_{i,h(i)},$$

where d is a prime, $\mathbb{F}_d$ is the finite field with d elements, and $k \ll d$. The degree of the gcd of every pair of monomials in $\mathsf{NW}_{d,k}$ is at most k. For concreteness, we fix $k = \lceil \sqrt{d} \rceil$. The family of polynomials $\mathcal{NW} := \{\mathsf{NW}_{d,k} : d \text{ is a prime}\}$ and close variants of it have been used as hard explicit polynomial families in several recent arithmetic circuit lower bound proofs. But, unlike the permanent, very little is known about the various structural and algorithmic/complexity aspects of $\mathcal{NW}$ beyond the fact that $\mathcal{NW} \in \mathsf{VNP}$. Is $\mathsf{NW}_{d,k}$ characterized by its symmetries? Is it circuit-testable, i.e., given a circuit C can we check efficiently if C computes $\mathsf{NW}_{d,k}$? What is the complexity of equivalence test for $\mathcal{NW}$, i.e., given black-box access to a $f \in \mathbb{F}[\mathbf{x}]$, can we check efficiently if there exists an invertible linear transformation A such that $f = \mathsf{NW}_{d,k}(A \cdot \mathbf{x})$? Characterization of polynomials by their symmetries plays a central role in the geometric complexity theory program. Here, we answer the first two questions and partially answer the third.

We show that $\mathsf{NW}_{d,k}$ is characterized by its group of symmetries over $\mathbb{C}$, but not over $\mathbb{R}$. We also show that $\mathsf{NW}_{d,k}$ is characterized by circuit identities which implies that $\mathsf{NW}_{d,k}$ is circuit-testable in randomized polynomial time. As another application of this characterization, we obtain the "flip theorem" for $\mathcal{NW}$.

We give an efficient equivalence test for $\mathcal{NW}$ in the case where the transformation A is a block-diagonal permutation-scaling matrix. The design of this algorithm is facilitated by an almost complete understanding of the group of symmetries of $\mathsf{NW}_{d,k}$: We show that if A is in the group of symmetries of $\mathsf{NW}_{d,k}$ then $A = D \cdot P$, where D and P are diagonal and permutation matrices respectively. This is proved by completely characterizing the Lie algebra of $\mathsf{NW}_{d,k}$, and using an interplay between the Hessian of $\mathsf{NW}_{d,k}$ and the evaluation dimension.

ISSN 1433-8092

1 Introduction

Proving super-polynomial lower bounds for Boolean and arithmetic circuits computing explicit functions is the holy grail of circuit complexity. Over the past few decades, research on lower bounds has gradually pushed the frontier by bringing in novel methods in the arena and carefully building upon the older ones. Some of the notable achievements are – lower bounds for AC^0 circuits [FSS81, Ajt83, Hås86], monotone circuits [Raz85, AB87], $\mathsf{ACC}(p)$ circuits [Raz87, Smo87] and ACC circuits [Wil14, MW18] in the Boolean case, and lower bounds for homogeneous depth three circuits [NW97], multilinear formulas [Raz09, RY09], homogeneous depth four circuits [GKKS14, KLSS17, KS17b] and the lower bound on the depth of circuits for MaxFlow [Mul99] in the arithmetic case. The slow progress in circuit lower bounds is explained by a few "barrier" type results, particularly by the notion of natural proofs [RR97] for Boolean circuits, and the notion of algebraically natural proofs [FSV17, GKSS17] for arithmetic circuits [1]. Most lower bound proofs, but not all [2], do fit in the natural proof framework.

It is apparent from the concept of natural proofs and its algebraic version that in order to avoid this barrier, we need to develop an approach that violates the so called *constructivity* criterion or the *largeness* criterion. Focusing on the latter criterion, it means, if an explicit function has a special property that random functions do not have, and if a lower bound proof for circuits computing this explicit function uses this special property critically, then such a proof circumvents the natural proof barrier automatically. For polynomial functions (simply polynomials), *characterization by symmetries* is such a special property[3], and the geometric complexity theory (GCT) program [MS01] is an approach to proving super-polynomial arithmetic circuit lower bound by crucially exploiting this property of the permanent and the determinant polynomials. From hereon, our discussion will be restricted to polynomial functions and arithmetic circuits.

The permanent family is complete for the class VNP and the determinant family is complete for the class VBP under p-projections. The class $\mathsf{VBP} \subseteq \mathsf{VP}$ consists of polynomial families that are computable by poly-size algebraic branching programs; this class has another interesting complete family, namely the iterated matrix multiplication (IMM) family. These three polynomial families have appeared in quite a few lower bound proofs [NW97, GK98, MR04, Raz09, RY09, GKKS14, FLMS15, KS17b, KST16b, CLS18] in the arithmetic circuit literature. That permanent and determinant are characterized by their respective groups of symmetries are classical results [MM62, Fro97]. It has also been shown that IMM is characterized by its symmetries [Ges16, KNST17]. There are two other polynomial families in VP, the power symmetric polynomials and the sum-product polynomials, that are known to possess this rare property (see Section 2 in [CKW11]). However, the elementary symmetric polynomial is not characterized by its symmetries [Hüt16].

In the recent years, another polynomial, namely the Nisan-Wigderson design polynomial (in

[1] Presently, the evidences in favor of existence of one-way functions (which implies the natural proof barrier) are much stronger than that of existence of succinct hitting-set generators (which implies the algebraically natural proof barrier). However, there are a few results in algebraic complexity that exhibit, unconditionally [EGdOW18] or based on more plausible complexity theoretic assumptions [BIJL18], the limitations of some of the current techniques in proving lower bounds for certain restricted arithmetic models.

[2] like the lower bounds for monotone and ACC circuits

[3] A random polynomial is not characterized by its symmetries with high probability (see Proposition 3.4.9 in [Gro12])

short, design polynomial), and close variants of it have been used intensely as hard explicit polynomials in several lower bound proofs for depth three, depth four and depth five circuits [KSS14, CM14, KS14a, KS14b, KLSS17, KS17b, KS16a, KS16c, KS16b, KST16a, FKS16, KS17a]. In some cases, the design polynomial (Definition 2.2) yielded lower bounds that are not known yet for the permanent, determinant and IMM (as in [KS16a, KS16b, FKS16, KS17a]). It can be easily shown that the design polynomial defines a family in VNP (see Observation B.1). But, very little is otherwise known about the various structural and algorithmic/complexity aspects of this family. Like the permanent, is it characterized by its symmetries? Is it circuit testable? What is the complexity of equivalence test for the Nisan-Wigderson design polynomial? It is reasonable to seek answers to these fundamental questions for a natural family like the design polynomials. Moreover, in the light of some recent developments in GCT [IP16, BIP16, IMW17], it may be worth studying other polynomial families (like the design polynomials and the IMM) that have some of the "nice features" of the permanent and the determinant and that may also fit in the GCT framework. We refer the reader to [Gro12, Aar17, Mul12, Reg02] for an overview of GCT. If the design polynomial family turns out to be in VP then that would be an interesting result by itself with potentially important complexity theoretic and algorithmic consequences. If a polynomial has a small depth-4 circuit, then it is a projection of a small NW design polynomial (see Observation B.2)

In this article, we answer some of the above questions on the design polynomial pertaining to its group of symmetries. Our results accord a fundamental status to this polynomial family.

1.1 Our results

Some of the basic definitions and notations are given in Section 2. The design polynomial $\mathsf{NW}_{d,k}$ is defined (in Definition 2.2) using two parameters, d (the degree) and k (the "intersection" parameter). Our results hold for any $k \in [1, \frac{d}{4} - 5]$, but (from the lower bound point of view) it is best to think of k as d^ϵ for some arbitrarily chosen constant $\epsilon \in (0, 1)$. The number of variables in $\mathsf{NW}_{d,k}$ is $n = d^2$. Any polynomial can be expressed as an affine projection of $\mathsf{NW}_{d,k}$, for a possibly large d (see Observation B.2). For notational convenience, we will drop the subscripts d and k whenever they are clear from the context. Let $\mathscr{G}_f$ be the group of symmetries of a polynomial f over an underlying field $\mathbb{F}$ (see Definition 2.6).

Theorem 1 (Characterization by symmetries)**.** *Let $\mathbb{F} = \mathbb{C}$ be the underlying field and f be a homogeneous degree-d polynomial in $n = d^2$ variables. If $\mathscr{G}_{\mathsf{NW}} \subseteq \mathscr{G}_f$ then $f = \alpha \cdot \mathsf{NW}$ for some $\alpha \in \mathbb{C}$.*

The theorem, proven in Section 3, holds over any field $\mathbb{F}$ having a d-th root of unity $\zeta \neq 1$ and $|\mathbb{F}| \neq d + 1$. We also show in Section 4.3 that NW is not characterized by its symmetries over $\mathbb{R}, \mathbb{Q}$ and finite fields not containing a d-th primitive root of unity – in contrast, the permanent is characterized by its symmetries over these fields. The symmetries of NW have a nice algorithmic application: Although, it is not known if NW is computable by a $\mathrm{poly}(d)$ size circuit (Definition 2.1), the following theorem shows that checking if a given circuit computes NW can be done efficiently. In this article, whenever we mention size-s circuit, we mean size-s circuit with degree bounded by $\delta(s)$, which is an arbitrarily fixed polynomial function[4] of s. Let $\mathbf{x}$ be the set of n variables of NW. We will identify a circuit with the polynomial computed by it.

[4] This is the interesting scenario in algebraic complexity theory as polynomial families in VP admit circuits with degree bounded by a polynomial function of size.

Theorem 2 (Circuit testability)**.** *There is a randomized algorithm that takes input as black-box access to a circuit* $\mathtt{C}(\mathbf{x})$ *of size* s *over a finite field* $\mathbb{F}$*, where* $|\mathbb{F}| \geq 4 \cdot \delta(s)$ *(recall* $\delta(s)$ *is an upper bound on the degree of size* s *circuits), and determines correctly whether or not* $\mathtt{C}(\mathbf{x}) = \mathsf{NW}$ *with high probability, using* $\mathrm{poly}(s)$ *field operations.*

A suitable version of the theorem also holds over $\mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$. Such a theorem is known for the permanent with two different proofs, one using self-reducibility of the permanent [Lip89] and the other using its symmetries [Mul10]. We do not know if NW has a self-reducible property like the permanent, but its symmetries are powerful enough to imply the above result. The theorem is proven in Section 5 by showing that NW is characterized by circuit identities over *any* field (see Definition 2.10). This characterization, which uses the symmetries of NW, also implies the following result. For this result, we can assume $\delta(s) \geq d$, without any loss of generality.

Theorem 3 (Flip theorem)**.** *Suppose* NW *is not computable by circuits of size* s *over a finite field* $\mathbb{F}$*, where* $|\mathbb{F}| \geq 4 \cdot \delta(s)$ *and* $\delta(s)$ *is an upper bound on the degree of size* s *circuits. Then, there exist points* $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{F}^n$*, where* $m = \mathrm{poly}(s)$*, such that for every circuit* $\mathtt{C}$ *over* $\mathbb{F}$ *of size at most* s*, there is an* $\ell \in [m]$ *satisfying* $\mathtt{C}(\mathbf{a}_\ell) \neq \mathsf{NW}(\mathbf{a}_\ell)$*. A set of randomly generated points* $\mathbf{a}_1, \ldots, \mathbf{a}_m \in_r \mathbb{F}^n$ *has this property with high probability. Moreover, black-box derandomization of polynomial identity testing for size-*$(10s)$ *circuits over* $\mathbb{F}$ *using* $\mathrm{poly}(s)$ *field operations implies that the above-mentioned points can be computed deterministically using* $\mathrm{poly}(s)$ *field operations.*

An appropriate version of the theorem also holds over $\mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$. The flip theorem is known for the permanent [Mul10, Mul11] [5]. Similar theorems have also been shown for the $3SAT$ problem [FPS08, Ats06]. Results of this kind show that if a certain function ($3SAT$ or permanent or NW) is not computable by small circuits then there exists a short list of efficiently computable "hard instances" that fail all small circuits.

We show another algorithmic application of the knowledge of the symmetries of NW in solving a natural case of the equivalence test problem for NW, namely block-diagonal permutation-scaling equivalence test (BD-PS equivalence test, in short). An equivalence test for NW checks if a given polynomial $f \in \mathbb{F}[\mathbf{x}]$ satisfies $f = \mathsf{NW}(A \cdot \mathbf{x})$, where A is an invertible linear transformation. A BD-PS equivalence test is the special case where A is a product of a block-diagonal permutation matrix and an invertible scaling matrix. The following theorem is proved in Section 6.

Theorem 4 (BD-PS equivalence test for NW)**.** *Let* $k \in [1, \frac{d}{3}]$*,* $\mathbb{F}$ *be a finite field such that* $d \nmid (|\mathbb{F}| - 1)$ *and* $|\mathbb{F}| \geq 4d$*. There is a randomized algorithm that takes input black-box access to a degree* d *polynomial* $f \in \mathbb{F}[\mathbf{x}]$ *and correctly decides if* f *is BD-PS equivalent to* NW *with high probability. If the answer is yes then it outputs a* A *such that* $f = \mathsf{NW}(A \cdot \mathbf{x})$*, where* A *is a product of a block-diagonal permutation matrix and an invertible scaling matrix. The running time is* $\mathrm{poly}(d, \log |\mathbb{F}|)$*.*

An appropriate version of the theorem holds over $\mathbb{R}$ (details given in Section F.4 of the Appendix). Efficient equivalence tests are known for the Permanent and IMM over $\mathbb{C}$, $\mathbb{Q}$ and finite fields [Kay12, KNST17] and for the Determinant over $\mathbb{C}$ and finite fields [Kay12, GGKS19]. In [Kay12], it was shown that equivalence test for the Permanent reduces to permutation-scaling (PS) equivalence test. We show in Section 6 that equivalence test for NW reduces to block-permuted equivalence test[6],i.e., we can assume without loss of generality that A is a block-permuted matrix. Theorem 4 solves the equivalence test for NW in the case where A is a block-diagonal matrix and

[5] We have borrowed the name 'flip theorem' from these work.

[6] It decides if there exists a block-permuted matrix (Definition 2.3) $A \in \mathrm{GL}_{d^2}(\mathbb{F})$ such that $f = \mathsf{NW}(A \cdot \mathbf{x})$

additionally has the permutation-scaling (PS) structure. Even this case is quite nontrivial and may serve as an important ingredient for an efficient general equivalence test for NW. The design of the test in Theorem 4 is facilitated by a near complete understanding of the symmetries of NW as stated in the following theorem. The proof is given in Section 4.2.

Theorem 5 (Structure of $\mathcal{G}_{\mathsf{NW}}$)**.** *Let $\mathbb{F}$ be the underlying field of size greater than $\binom{d}{2}$ and $\mathrm{char}(\mathbb{F}) \neq d$. If $A \in \mathcal{G}_{\mathsf{NW}}$ then $A = D \cdot P$, where $D, P \in \mathcal{G}_{\mathsf{NW}}$ are diagonal and permutation matrices respectively.*

The group of symmetries of the permanent has a similar structure [MM62]. The above structure also plays a crucial role in showing that NW is not characterized by its symmetries over $\mathbb{R}$. The proof of the theorem involves a complete characterization of the Lie algebra of NW, and an interplay between the Hessian of NW and the evaluation dimension measure. We first prove the structural results (Theorems 1 and 5) and then show their algorithmic applications (Theorems 2, 3 and 4). Because of space constraint, the proof details of the theorems are shifted to the appendix. A comparison between the Permanent and NW is summarized in a table in Section A.

2 Preliminaries

Notations. The set of natural numbers is $\mathbb{N} = \{0, 1, 2 \ldots\}$ and $\mathbb{N}^{\times} = \mathbb{N} \backslash \{0\}$. For $r \in \mathbb{N}^{\times}$, $[r] = \{0, \ldots, r-1\}$. The general linear group $\mathrm{GL}_r(\mathbb{F})$ is the group of all $r \times r$ invertible matrices over $\mathbb{F}$. Throughout this article, $\mathrm{poly}(r)$ means $r^{O(1)}$ and $\exp(r)$ means 2^r. For a prime d, $\mathbb{F}_d$ is the finite field of order d whose elements are naturally identified with $[d] = \{0, 1, \ldots, d-1\}$. Let $\mathbf{x}$ be the following disjoint union of variables,

$$\mathbf{x} := \biguplus_{i \in [d]} \mathbf{x}_i, \tag{1}$$

where $\mathbf{x}_i := \{x_{i,0}, \ldots, x_{i,d-1}\}$. The total number of variables in $\mathbf{x}$ is $n = d^2$. $\mathbb{F}[\mathbf{x}]$ and $\mathbb{F}_d[z]$ denote the rings of multivariate and univariate polynomials over $\mathbb{F}$ and $\mathbb{F}_d$ in $\mathbf{x}$ and z variables respectively, and the set $\mathbb{F}_d[z]_k := \{h \in \mathbb{F}_d[z] : \deg(h) \leq k\}$. We will represent elements of $\mathbb{F}$ by lower case Greek alphabets ($\alpha, \beta, \ldots$), elements of $\mathbb{F}_d$ by lower case Roman alphabets ($a, b, \ldots$), multivariate polynomials over $\mathbb{F}$ by f, g and q, univariate polynomials over $\mathbb{F}_d$ by p and h, matrices over $\mathbb{F}$ by capital letters ($A, B, C, \ldots$), and the set of variables by $\mathbf{x}, \mathbf{y}, \mathbf{z}$ and vectors over $\mathbb{F}$ by $\mathbf{a}, \mathbf{b}$. Variable sets are interpreted as column vectors when left multiplied to a matrix. For instance, in $A \cdot \mathbf{x}$, $\mathbf{x}$ is the vector $(x_{0,0}\ x_{0,1} \ldots\ x_{0,d-1}\ \ldots\ x_{d-1,0}\ x_{d-1,1} \ldots\ x_{d-1,d-1})^T$, and we say A is applied on $\mathbf{x}$.

2.1 Algebraic preliminaries

A polynomial f is homogeneous if the degree of all the monomials of f are the same. Polynomial $f \in \mathbb{F}[\mathbf{x}]$ is set-multilinear in the sets $\mathbf{x}_0, \ldots, \mathbf{x}_{d-1}$ (as defined in Equation (1)) if every monomial contains exactly one variable from each set $\mathbf{x}_i$ for $i \in [d]$.

Definition 2.1 (Arithmetic circuit)**.** An arithmetic circuit $\mathtt{C}$ over $\mathbb{F}$ is a directed acyclic graph in which a node with in-degree zero is labelled with either a variable or a $\mathbb{F}$-element, an edge is labelled with a $\mathbb{F}$-element, and other nodes are labelled with $+$ and $\times$. Computation proceeds in a natural way: a node with in-degree zero computes its label, an edge scales a polynomial by its label, and a node labelled with $+/\times$ computes the sum/product of the polynomials computed

at the end of the edges entering the node. The polynomials computed by nodes with out-degree zero are the output of $\mathtt{C}$. The *size* of $\mathtt{C}$ is the sum of the number of nodes and edges in the graph. The *degree* of $\mathtt{C}$ is the maximum over the degree of the polynomials computed at all nodes of $\mathtt{C}$.

Definition 2.2 (Nisan-Wigderson polynomial)**.** Let $d > 2$ be a prime and $k \in \mathbb{N}$. The Nisan-Wigderson design polynomial is defined as in [KSS14] (which is inspired by the Nisan-Wigderson set-systems [NW94]),

$$\mathsf{NW}_{d,k}(\mathbf{x}) := \sum_{h \in \mathbb{F}_d[z]_k} \prod_{i \in \mathbb{F}_d} x_{i,h(i)}.$$

It is a degree-d homogeneous and set-multilinear polynomial in $n = d^2$ variables, having d^{k+1} monomials. We drop the subscripts d, k for notational convenience. NW satisfies the 'low intersection' property, meaning any two monomials of NW have at most k variables in common. This follows from the fact that the monomials are obtained from polynomials in $\mathbb{F}_d[z]_k$.

Definition 2.3 (Block-permuted matrix)**.** A matrix $A \in \mathbb{F}^{d^2 \times d^2}$ is a block-permuted matrix with block size d if $A = B \cdot (P \otimes I_d)$, where $B \in \mathbb{F}^{d^2 \times d^2}$ is a block-diagonal matrix with block size d, $P \in \mathbb{F}^{d \times d}$ is a permutation matrix, and I_d is the $d \times d$ identity matrix.

Definition 2.4 (Evaluation dimension)**.** Let $f \in \mathbb{F}[\mathbf{y}]$ and $\mathbf{z} \subseteq \mathbf{y}$. The evaluation dimension of f with respect to $\mathbf{z}$ is,

$$\text{evalDim}_{\mathbf{z}}(f) := \dim(\ \mathbb{F}\text{-span}\ \{f(\mathbf{y})|_{\mathbf{z}=\mathbf{a}} : \mathbf{a} \in \mathbb{F}^{|\mathbf{z}|}\}\ ).$$

Definition 2.5 (Hessian)**.** Let $f \in \mathbb{F}[\mathbf{y}]$ be a polynomial in $\mathbf{y} = \{y_1, y_2, \ldots, y_n\}$ variables. The Hessian of f is the following matrix in $(\mathbb{F}[\mathbf{y}])^{n \times n}$,

$$H_f(\mathbf{y}) := \left(\frac{\partial^2 f}{\partial y_i \cdot \partial y_j}\right)_{i,j \in [n]}.$$

We would need the following property of $H_f(\mathbf{y})$ that can be proved using chain-rule of derivatives.

Lemma 2.1 (Lemma 2.6 of [CKW11])**.** *Let $g \in \mathbb{F}[\mathbf{y}]$ and $f = g(A \cdot \mathbf{y})$ for some $A \in \mathbb{F}^{n \times n}$. Then,*

$$H_f(\mathbf{y}) = A^T \cdot H_g(A \cdot \mathbf{y}) \cdot A.$$

Definition 2.6 (Group of symmetries)**.** Let $f \in \mathbb{F}[\mathbf{y}]$ be an n-variate polynomial. The set $\mathscr{G}_f = \{A \in \mathrm{GL}_n(\mathbb{F}) : f(A \cdot \mathbf{y}) = f(\mathbf{y})\}$ forms a group under matrix multiplication and it is called the group of symmetries of f over $\mathbb{F}$.

Definition 2.7 (Lie algebra)**.** Let $f \in \mathbb{F}[\mathbf{y}]$ be a polynomial in $\mathbf{y} = \{y_1, y_2, \ldots, y_n\}$ variables. The Lie algebra of f, denoted by $\mathfrak{g}_f$, is the set of matrices $B = (b_{i,j})_{i,j \in [n]} \in \mathbb{F}^{n \times n}$ satisfying the relation,

$$\sum_{i,j \in [n]} b_{i,j} \cdot y_j \cdot \frac{\partial f}{\partial y_i} = 0.$$

It is easy to check that $\mathfrak{g}_f$ is a vector space over $\mathbb{F}$. The following property relates the Lie algebras of $f(\mathbf{y})$ and $f(A \cdot \mathbf{y})$ for $A \in \mathrm{GL}_n(\mathbb{F})$. See Proposition 58 of [Kay12] for a proof of this fact.

Lemma 2.2 (Conjugacy of Lie algebras)**.** *Let $g \in \mathbb{F}[\mathbf{y}]$ be an n-variate polynomial. If $f(\mathbf{y}) = g(A \cdot \mathbf{y})$ for $A \in GL_n(\mathbb{F})$, then $\mathfrak{g}_f = A^{-1} \cdot \mathfrak{g}_g \cdot A$.*

Lemma 2.3. [Kay12] *Given black-box access to an n-variate degree d polynomial $f \in \mathbb{F}[\mathbf{x}]$, a basis of $\mathfrak{g}_f$ can be computed in randomized $poly(n, d, \rho)$ time, where ρ is the bit complexity of the coefficients of f.*

Over $\mathbb{C}$, the Lie algebra $\mathfrak{g}_f$ is related to the group of symmetries $\mathscr{G}_f$ as stated in the following definition. For $B \in \mathbb{C}^{n\times n}$, let $e^B := \sum_{i\in\mathbb{N}} \frac{B^i}{i!} \in \mathbb{C}^{n\times n}$ (the series always converges).

Definition 2.8 (Continuous and discrete symmetries)**.** Let $f \in \mathbb{C}[\mathbf{y}]$. If $A \in \mathfrak{g}_f$ then $e^{tA} \in \mathscr{G}_f$ for every $t \in \mathbb{R}$ (see [Hal15] for a proof of this fact). Elements of the set $\{e^{tA} : A \in \mathfrak{g}_f \text{ and } t \in \mathbb{R}\}$ are the continuous symmetries of f. All the other symmetries in $\mathscr{G}_f$ are the discrete symmetries of f.

Definition 2.9 (Characterization by symmetries)**.** A homogeneous degree-d polynomial $g \in \mathbb{F}[\mathbf{y}]$ is said to be *characterized by its symmetries* if for every degree-d homogeneous polynomial $f \in \mathbb{F}[\mathbf{y}]$, $\mathscr{G}_g \subseteq \mathscr{G}_f$ implies that $f(\mathbf{y}) = \alpha \cdot g(\mathbf{y})$ for some $\alpha \in \mathbb{F}$.

Definition 2.10 (Characterization by circuit identities)**.** Let $g \in \mathbb{F}[\mathbf{y}]$ be an n-variate polynomial, and $\mathbf{z}, \mathbf{u}$ be two sets of constantly many variables and $|\mathbf{z}| = c$. Suppose that there exist $m = \text{poly}(n)$ polynomials $q_1(\mathbf{z}, \mathbf{u}), \ldots, q_m(\mathbf{z}, \mathbf{u})$ over $\mathbb{F}$ such that for every $i \in [m]$, q_i is computable by a constant size circuit and there are matrices $A_{i1}, \ldots, A_{ic} \in \mathbb{F}[\mathbf{u}]^{n\times n}$ computable by $\text{poly}(n)$ size circuits, and the following condition is satisfied: For $f \in \mathbb{F}[\mathbf{y}]$, $q_i(f(A_{i1} \cdot \mathbf{y}), \ldots, f(A_{ic} \cdot \mathbf{y}), \mathbf{u}) = 0$ for every $i \in [m]$ if and only if $f = \alpha \cdot g$ for some $\alpha \in \mathbb{F}$. Then, g is *characterized by circuit identities* over $\mathbb{F}$.

The above definition is taken (after slight modifications to suit our purpose) from Definition 3.4.7 in [Gro12] and is attributed to an article by Mulmuley [Mul07].

3 Characterization of NW by symmetries and circuit identities

3.1 Symmetry characterization: Theorem 1

Let $\mathbb{F}$ be a field having a d-th root of unity $\zeta \neq 1$ and $|\mathbb{F}| \neq d + 1$.[7] As d is a prime, ζ is primitive, i.e., $\zeta^d = 1$ and $\zeta^t \neq 1$ for $0 < t < d$. The rows and columns of a matrix in $\mathscr{G}_{\mathsf{NW}}$ are indexed by the set $\{(i, j) : i, j \in \mathbb{F}_d\}$.

Claim 3.1. *The following matrices in $\mathbb{F}^{n\times n}$ are in $\mathscr{G}_{\mathsf{NW}}$:*

1. *A_β, a diagonal matrix with $((i, j), (i, j))$-th entry as $\beta_i \in \mathbb{F}^\times$ for $i, j \in [d]$, such that $\prod_{i\in[d]} \beta_i = 1$.*
2. *A_ℓ, a diagonal matrix with $((i, j), (i, j))$-th entry as $\zeta^{i^\ell \cdot j}$ for $i, j \in [d]$, where $\ell \in [d - k - 1]$.* [8]
3. *A_h, $h \in \mathbb{F}_d[z]_k$, the $((i, j), (i, j + h(i)))$-th entry of A_h is 1 for $i, j \in [d]$ and other entries are 0.*

The proof of Claim 3.1 is given in Section C.1. The matrices A_β are the continuous symmetries while A_ℓ, A_h are discrete symmetries of NW for all choices of $\boldsymbol{\beta}, \ell, h$. The symmetries in 2 are very different from the symmetries of the Determinant and the Permanent. The following Claim immediately implies Theorem 1. Its proof is given in Section C.2.

Claim 3.2. *Let f be a homogeneous degree-d polynomial in $\mathbb{F}[\mathbf{x}]$. If $\mathscr{G}_f$ contains the matrices A_β, A_ℓ and A_h (for all choices of $\boldsymbol{\beta}$, ℓ and h, as mentioned in Claim 3.1) then $f = \alpha \cdot \mathsf{NW}$ for some $\alpha \in \mathbb{F}$.*

[7] For a prime d, $|\mathbb{F}| = d + 1$ if and only if d is a Mersenne prime.
[8] Recall, $[d - k - 1] = \{0, 1, \ldots, d - k - 2\}$

3.2 Characterization by circuit identities

In the following lemma, we show that NW is characterized by circuit identities (as defined in Definition 2.10). The lemma is crucially used in the proofs of Theorems 2 and 3 in Section 5. Its proof is given in Section C.3.

Lemma 3.1. *Polynomial* NW *is characterized by circuit identities over any field* $\mathbb{F}$.

4 Lie algebra and symmetries of NW

We first give a complete description of the Lie algebra of NW by giving an explicit $\mathbb{F}$-basis. Then, using this knowledge, we analyse the structure of the symmetries of NW and prove Theorem 5. Thereafter, using Theorem 5, we show that NW is not characterized by its symmetries over fields that do not contain a d-th primitive root of unity. The rows and columns of a $n \times n$ matrix in $\mathfrak{g}_{\mathsf{NW}}$ and $\mathscr{G}_{\mathsf{NW}}$ are indexed by the set $\{(i,j) : i,j \in \mathbb{F}_d\}$, which is naturally identified with the $\mathbf{x}$-variables, where $\mathbf{x} = (x_{0,0} \dots x_{0,d-1} \ \dots \ x_{d-1,0} \dots x_{d-1,d-1})^T$.

4.1 Lie algebra of NW

It turns out that the Lie algebra of NW is a subspace of the Lie algebra of every set-multilinear polynomial. (The default partition of a set-multilinear polynomial is $\mathbf{x} = \uplus_{i\in[d]}\mathbf{x}_i$.)

Lemma 4.1. *Let* $\mathbb{F}$ *be a field and* $\mathrm{char}(\mathbb{F}) \neq d$. *The dimension of* $\mathfrak{g}_{\mathsf{NW}}$ *over* $\mathbb{F}$ *is* $d-1$, *and the diagonal matrices* $B_1, \dots, B_\ell$ *(defined below) form a* $\mathbb{F}$*-basis of* $\mathfrak{g}_{\mathsf{NW}}$. *For* $\ell \in \{1, \dots, d-1\}$,

$$(B_\ell)_{(i,j),(i,j)} = \begin{cases} 1, & \text{if } i = 0, j \in [d] \\ -1, & \text{if } i = \ell, j \in [d] \\ 0, & \text{otherwise.} \end{cases}$$

The lemma is proven in Section D.1 by carefully analysing a system of linear equations obtained from the monomials of NW. It follows that every $B \in \mathfrak{g}_{\mathsf{NW}}$ is of the form $\mathrm{diag}(\alpha_0, \dots, \alpha_{d-1}) \otimes I_d$, where each $\alpha_i \in \mathbb{F}$ and $\sum_{i\in[d]} \alpha_i = 0$. It follows that the continuous symmetries of NW consist of matrices of the form $A = \mathrm{diag}(\beta_0, \dots, \beta_{d-1}) \otimes I_d$, where each $\beta_i \in \mathbb{C}$ and $\prod_{i\in[d]} \beta_i = 1$.

4.2 Structure of $\mathscr{G}_{\mathsf{NW}}$: Theorem 5

Lemma 4.1 implies the following.

Claim 4.1. *Every* $A \in \mathscr{G}_{\mathsf{NW}}$ *is a block-permuted matrix with block size* d.

The proof of the claim is given in Section D.2. Using Claim 4.1, Hessian and the evaluation dimension of NW, we give a proof of Theorem 5 in Section D.3.

4.3 NW is not characterized by its symmetries over $\mathbb{R}$

Let $\mathbb{F}$ be either $\mathbb{R}, \mathbb{Q}$ or a finite field such that $d \nmid |\mathbb{F}| - 1$. Then, $\mathbb{F}$ does not contain a d-th primitive root of unity, and so the matrices A_ℓ, for $\ell \in [d-k-1]$ mentioned in Claim 3.1, are no longer the

symmetries of NW over $\mathbb{F}$. The next lemma shows that over such $\mathbb{F}$ all the diagonal symmetries of NW are of the type A_β mentioned in Claim 3.1. This then implies the following theorem, which may seem somewhat surprising as we do not know all the permutation symmetries of NW. The proofs are given in Section D.4.

Lemma 4.2. *If $D \in \mathscr{G}_{\mathsf{NW}}$ is a diagonal matrix over $\mathbb{F}$ then $D = diag(\beta_0, \ldots, \beta_{d-1}) \otimes I_d$, where each $\beta_i \in \mathbb{F}$ and $\prod_{i\in[d]} \beta_i = 1$.*

Theorem 6. NW *is not characterized by its symmetries over* $\mathbb{F}$.

5 Circuit testability and the flip theorem for NW

In this and the next section, we show that the knowledge of the symmetries of NW plays a crucial role in answering some of the algorithmic questions related to NW. This section is devoted to Theorems 2 and 3. The main ingredient of their proofs is Lemma 3.1. We present the circuit testing algorithm here and push the proof of the Flip theorem to Section E.

Proof of Theorem 2:
Let $\mathtt{C}$ be a given circuit of size s over $\mathbb{F}$ that computes an n-variate polynomial $f = \mathtt{C}(\mathbf{x})$. Naturally, $\deg(f) \leq \delta(s)$. Algorithm 1 intends to check, in steps 2 and 3, if f satisfies the identities given in the proof of Lemma 3.1. If $f \neq \alpha \cdot \mathsf{NW}$ for all $\alpha \in \mathbb{F}$, then at least one of the identities is not satisfied. For the polynomials q_1, q_2 and q_3 defined in the proof of Lemma 3.1, observe that the degree of $q_1(f(A_i(u)\cdot\mathbf{x}), f(\mathbf{x}), u)$ is bounded by $2\cdot\delta(s)$, whereas the degrees of $q_2(f(A_{a,r}\cdot\mathbf{x}), f(\mathbf{x}))$ and $q_3(f(A_t\cdot\mathbf{x}))$ are at most $\delta(s)$. As $|\mathbb{F}| \geq 4\cdot\delta(s)$, by Schwartz-Zippel lemma [Zip79,Sch80], step 4 returns 'False' with probability at least $\frac{1}{2}$. If $f = \alpha\cdot\mathsf{NW}$ for some $\alpha \in \mathbb{F}$ then all the identities are satisfied, and step 7 ensures that $\alpha = 1$. Clearly, the algorithm uses $\mathrm{poly}(s)$ field operations. The success probability is boosted from $\frac{1}{2}$ to $1 - \exp(-s)$ by repeating the algorithm $\mathrm{poly}(s)$ times.

Algorithm 1 Circuit testing for NW

Input: Black-box access to a circuit $\mathtt{C}$ of size s over $\mathbb{F}$.
Output: 'True' if $\mathtt{C}(\mathbf{x}) = \mathsf{NW}$, else 'False'.

1. Pick $\mathbf{a} \in_r \mathbb{F}^n$ and $\mu \in_r \mathbb{F}$.
2. **for** $i \in [d], a \in \mathbb{F}_d^\times, r \in [k+1], t \in [d]\backslash[k+1]$ **do**
3. **if** $(\mathtt{C}(A_i(\mu)\cdot\mathbf{a}) - \mu\cdot\mathtt{C}(\mathbf{a}) \neq 0)$ or $(\mathtt{C}(A_{a,r}\cdot\mathbf{a}) - \mathtt{C}(\mathbf{a}) \neq 0)$ or $(\mathtt{C}(A_t\cdot\mathbf{a}) \neq 0)$ **then**
4. **return** 'False'.
5. **end if**
6. **end for**
7. Let $\mathbf{b} \in \mathbb{F}^n$ be an assignment obtained by setting $x_{i0} = 1$, for $i \in [d]$, and all other variables to zero. If $f(\mathbf{b}) \neq 1$, return 'False'. Else, return 'True'.

6 Equivalence test for NW

First, we show a randomized reduction of equivalence test for NW to block-permuted equivalence test (in short, BP equivalence test) in Lemma 6.1. Then, we give an efficient equivalence test for

NW in the special case where the linear transformation is block-diagonal and is a product of a permutation matrix and a scaling matrix (Theorem 4).

Lemma 6.1 (Reduction to BP equivalence test). *Let $\mathbb{F}$ be a field such that $\mathrm{char}(\mathbb{F}) \neq d$ and $|\mathbb{F}| \geq 2d^2$. There is a randomized algorithm that takes input as black-box access to a degree d polynomial $f \in \mathbb{F}[\mathbf{x}]$ and does the following with high probability: It outputs black-box access to a degree d polynomial $g \in \mathbb{F}[\mathbf{x}]$ such that f is equivalent to NW if and only if g is BP equivalent to NW. Moreover, the transformation for f can be recovered efficiently from the transformation for g. The running time of this reduction is $\mathrm{poly}(d, \rho)$, where ρ is the bit complexity of the coefficients of f* [9].

Algorithm 2 Reduction of equivalence test for NW to BP equivalence test

Input: Black-box access to $f \in \mathbb{F}[\mathbf{x}]$.
Output: Black-box access to $g \in \mathbb{F}[\mathbf{x}]$.

1. Compute a basis $L_1, \ldots, L_r$ of $\mathfrak{g}_f$. If $r \neq d-1$, output 'f is not equivalent to NW'.
2. Let S be an arbitrary subset of $\mathbb{F}$ of size d^2. Let $L = a_1 L_1 + \ldots + a_r L_r$, where $a_i \in_r S$. Compute $D \in \mathrm{GL}_{d^2}(\mathbb{F})$ such that $D^{-1} \cdot L \cdot D = \mathrm{diag}(\beta_1, \ldots, \beta_d) \otimes I_d$, where $\beta_j \in \mathbb{F}$. If no such D exists then output 'f is not equivalent to NW.'
3. Output black-box access to $f(D \cdot \mathbf{x})$.

Proof of correctness: The efficiency of Step 1 follows from Lemma 2.3. The correctness of Step 2 and 3 follow from the next claim whose proof is given in Section F.1.

Claim 6.1. *With high probability, matrix D can be computed in $\mathrm{poly}(d, \rho)$ time. Moreover, f is equivalent to NW if and only if $f(D \cdot \mathbf{x})$ is BP equivalent to NW.*

6.1 BD-PS equivalence test for NW: Theorem 4

Lemma 6.1 implies that to solve equivalence test for NW it is sufficient to focus on BP equivalence test. Here, we solve a special case of BP equivalence test, namely BD-PS equivalence test. We prove Theorem 4 in two steps: first we reduce BD-PS equivalence test to scaling equivalence test and then solve the scaling equivalence test. The algorithm pretends that f is BD-PS equivalent to NW and computes a block-diagonal permutation matrix A and an invertible scaling matrix B. In the end, the circuit testing algorithm of NW (Algorithm 1) is used to check if $f(A^{-1} \cdot B^{-1} \cdot \mathbf{x}) = \mathrm{NW}$.

6.1.1 Reduction of BD-PS equivalence test to scaling equivalence test

Assume $f = \mathrm{NW}(B \cdot A \cdot \mathbf{x})$, where A is a block-diagonal permutation matrix and B is an invertible scaling matrix. Algorithm 3 does not explicitly use the knowledge of the entries of B. Thus, we may assume without loss of generality that $B = I_{d^2}$. Then, the task reduces to solving the BD permutation equivalence test for NW. We identify matrix A with d permutations $\sigma_0, \ldots, \sigma_{d-1}$ on $[d]$ as $A = \mathrm{diag}(M_{\sigma_0}, \ldots, M_{\sigma_{d-1}})$, where M_{σ_i} is the $d \times d$ permutation matrix corresponding to σ_i [10].

Observation 6.1. *Suppose f is BD permutation equivalent to NW, i.e. $f = \mathrm{NW}(A \cdot \mathbf{x})$. Then, a monomial $\prod_{i \in \mathbb{F}_d} x_{i,h(i)}$ of NW gets mapped to a unique monomial $\prod_{i \in \mathbb{F}_d} x_{i,\sigma_i(h(i))}$ of f.*

[9] We assume that univariate polynomial factorization over $\mathbb{F}$ can be done in polynomial time.
[10] For $i, r, s \in [d]$, $M_{\sigma_i}(r, s) = 1$ if and only if $\sigma_i(r) = s$.

Algorithm 3 starts by assuming that $\sigma_0(0) = \cdots = \sigma_k(0) = 0$ and $\sigma_0(1) = 1$. The symmetries of NW allow us to make this assumption without loss of generality (Claim 6.2). The aim is to figure out all the entries of σ_i [11]. This is done by carefully picking a bunch of polynomials from $\mathbb{F}_d[z]_k$ (which we call *nice* polynomials) and then exploiting the association between f and NW mentioned in Observation 6.1 using these polynomials. The algorithm works over every field.

Algorithm 3 Block-diagonal permutation equivalence test for *NW*

Input: Black-box access to $f \in \mathbb{F}[\mathbf{x}]$.
Output: Black-box access to $g \in \mathbb{F}[\mathbf{x}]$ such that if f is BD-PS equivalent to NW then g is scaling equivalent to NW.

1. Assume that $\sigma_0(0) = \cdots = \sigma_k(0) = 0$ and $\sigma_0(1) = 1$ (Claim 6.2).
2. Construct a list of nice polynomials in $\mathbb{F}_d[z]_k$ (Definition 6.1) as mentioned in Claim 6.3.
3. Recover $(d-k)$ distinct entries of each permutation $\sigma_0, \ldots, \sigma_{d-1}$ as mentioned in Claim 6.4.
4. Let N be a $d \times d$ matrix, where the columns and rows are indexed by $(\sigma_0, \ldots, \sigma_{d-1})$ and $(0, \ldots, d-1)$ respectively and for $l, i \in [d], N(l,i) := \sigma_i(l)$. Pick $l_0, \ldots, l_k \in [d]$ such that in each of the rows indexed by $l_0, \ldots, l_k$ at least $k+1$ entries are known (Claim 6.5).
5. Use $l_0, \ldots, l_k \in [d]$ to recover all the entries of the rows of N as mentioned in Claim 6.6. Compute $A = \mathrm{diag}(M_{\sigma_0}, \ldots, M_{\sigma_{d-1}})$ and return black box access to $f(A^{-1} \cdot \mathbf{x})$

Proof of correctness: The following chain of claims argue the correctness of the algorithm. Their proofs are given in Section F.2. In these claims, ρ is the bit complexity of the coefficients of f.

Claim 6.2. (Canonical form of $\sigma_0, \ldots, \sigma_{d-1}$)*: Suppose $f \in \mathbb{F}[\mathbf{x}]$ is BD permutation equivalent to* NW*. Then, there exist permutations $\sigma_0, \ldots, \sigma_{d-1}$ on $[d]$ such that $\sigma_0(0) = \cdots = \sigma_k(0) = 0, \sigma_0(1) = 1$ and $A = \mathrm{diag}(M_{\sigma_0}, \ldots, M_{\sigma_{d-1}})$ satisfies $f = \mathrm{NW}(A \cdot \mathbf{x})$.*

Definition 6.1. (List of nice polynomials in $\mathbb{F}_d[z]_k$): $\{h_0, \ldots, h_{d-k-1}\} \subseteq \mathbb{F}_d[z]_k$ is called a list of nice polynomials if the following properties are satisfied:

1. For distinct $r_1, r_2 \in [d-k]$, $h_{r_1}(\ell) = h_{r_2}(\ell)$ for every $\ell \in [k]$ and $h_{r_1}(\ell) \neq h_{r_2}(\ell)$ for every $\ell \in \{k, \ldots, d-1\}$.
2. For every $r \in [d-k], \sigma_0(h_r(0)), \ldots, \sigma_k(h_r(k))$ can be computed in $\mathrm{poly}(d, \rho)$ time.

Claim 6.3. *A list of $d-k$ nice polynomials $\{h_0, \ldots, h_{d-k-1}\}$ can be computed in $\mathrm{poly}(d, \rho)$ time.*

Using the list of nice polynomials, we recover $d-k$ distinct entries of $\sigma_0, \ldots, \sigma_{d-1}$.

Claim 6.4. *Given a list of nice polynomials $\{h_0, \ldots, h_{d-k-1}\}$, we can recover $d-k$ distinct entries in each of $\sigma_0, \ldots, \sigma_{d-1}$ in $\mathrm{poly}(d, \rho)$ time.*

The matrix N defined in the algorithm is filled with some known entries and some unknowns. The goal is to recover all the entries of N which is accomplished by the following claims.

Claim 6.5. *Suppose $k \in [1, \frac{d}{3}]$. Then, there exist $k+1$ rows in N such that in each of these rows at least $k+1$ entries are known.*

Claim 6.6. *Using $k+1$ rows of N indexed by $l_0, \ldots, l_k$ (as mentioned in Step 4), we can recover all the entries of N in $\mathrm{poly}(d, \rho)$ time.*

[11] σ_i is treated as an ordered tuple $(\sigma_i(0), \ldots, \sigma_i(d-1))$

6.1.2 Scaling equivalence test for NW

We present an algorithm for solving the scaling equivalence test for NW over a finite field $\mathbb{F}$, where $d \nmid |\mathbb{F}| - 1$. The same algorithm with appropriate modifications works over $\mathbb{R}$. More details on this are given in Section F.4. Assume that f is scaling equivalent to NW.

Algorithm 4 Scaling equivalence test for *NW* over finite fields

Input: Black box access to $f \in \mathbb{F}[\mathbf{x}]$.
Output: An invertible diagonal matrix B such that $f = \mathsf{NW}(B \cdot \mathbf{x})$.

1. Let $B = \mathrm{diag}(\alpha_{0,0}, \ldots, \alpha_{d-1,d-1})$, where $\{\alpha_{i,j} : i,j \in [d]\}$ are unknown. Set $\alpha_{1,0} = \ldots = \alpha_{d-1,0} = 1$ (Claim 6.7).
2. Let $S = (0, z, \ldots, (d-1)z,\ 1, z+1 \ldots, (d-1)z+1,\ \ldots\ , d-2, z+d-2 \ldots, (d-1)z+d-2,\ d-1)$ be the ordered set of $d^2 - d + 1$ polynomials in $\mathbb{F}[z]$. For every $h \in S$, query the coefficient c_h of the monomial $\prod_{i \in \mathbb{F}_d} x_{i,h(i)}$ from the black-box of f (Observation 6.2).
3. Let C be a $0/1$ matrix of size $(d^2 - d + 1) \times (d^2 - d + 1)$ whose rows and columns are indexed by S and $\mathbf{y} = (y_{0,0}, \ldots, y_{0,d-1}, y_{1,1}, \ldots, y_{1,d-1}, \ldots, y_{d-1,1}, \ldots, y_{d-1,d-1})$, respectively, such that for $h \in S$ and $y_{i,j} \in \mathbf{y}$, the $(h, y_{i,j})$-th entry of C is 1 if $h(i) = j$. (It is argued in Claim D.3 that $|\det(C)|$ is a power of d). Compute the inverse of $\det(C)$ in $\mathbb{Z}_{|\mathbb{F}|-1}$ and denote it by γ. (Note that $\mathbf{y}$ does not contain the variables $\{y_{1,0}, \ldots, y_{d-1,0}\}$.)
4. Fix $\alpha_{i,j} \in \{\alpha_{0,0}, \ldots, \alpha_{d-1,d-1}\} \setminus \{\alpha_{1,0}, \ldots, \alpha_{d-1,0}\}$ arbitrarily. For every $h \in S$, compute the minor of C with respect to the row and column indexed by h and $y_{i,j}$ respectively and call it δ_h. Set $\alpha_{i,j} = \prod_{h \in S} c_h^{(\delta_h \cdot \gamma) \mod (|\mathbb{F}|-1)}$.
5. Set $B = \mathrm{diag}(\alpha_{0,0}, \ldots, \alpha_{d-1,d-1})$. Return B. (see Claim 6.8)

Proof of correctness: The following claims and observations argue the correctness of the algorithm. The proofs of the claims are given in Section F.3.

Claim 6.7. *We can assume that $\alpha_{1,0} = \ldots = \alpha_{d-1,0} = 1$ without loss of generality.*

The following observation can be proved easily.

Observation 6.2. *Given a monomial m, we can recover the coefficient of m in f in $\mathrm{poly}(d, \rho)$ time.*

Claim 6.8. *In Step 4, $\alpha_{i,j}$ can be computed in $\mathrm{poly}(d, \rho)$ time. Further, $f = \mathsf{NW}(B \cdot \mathbf{x})$.*

7 Few problems

In conclusion, we state a few problems on the NW polynomial which, if resolved, would shed more light on this fundamental polynomial family.

1. Is the $\mathcal{NW} = \{\mathsf{NW}_{d,k} : d \text{ is a prime}\}$ family VNP-complete for a suitable choice of k (say, $k = d^\epsilon$ for a constant $\epsilon > 0$)?
2. Is there an efficient algorithm to check if $\mathsf{NW}(\mathbf{a}) = 0$ at a given point $\mathbf{a} \in \{0,1\}^n$? This problem was also posed in [BS07] [12].

[12] We thank Andrej Bogdanov for pointing this out to us.

3. Is there an efficient general equivalence test for NW? Theorem 4 may turn out to be a vital ingredient in such a test.

4. Give a complete description of the permutation symmetries of NW. Are all the permutation symmetries captured in Lemma G.1 mentioned in Section G?

For the permanent polynomial, the solutions to these problems are well known.

Acknowledgment

We would like to thank Neeraj Kayal, Meena Mahajan for some insightful discussions on the design polynomial family. NG would also like to thank Anuj Tawari for his time in sitting through a few presentations on the proof of Theorem 4.

References

[Aar17] Scott Aaronson. P=?NP. *Electronic Colloquium on Computational Complexity (ECCC)*, 24:4, 2017.

[AB87] Noga Alon and Ravi B. Boppana. The monotone circuit complexity of boolean functions. *Combinatorica*, 7(1):1–22, 1987.

[Ajt83] Miklós Ajtai. Σ_1^1-formulae on finite structures. *Annals of Pure and Applied Logic*, 24(1):1–48, 1983.

[Ats06] Albert Atserias. Distinguishing SAT from polynomial-size circuits, through black-box queries. In *21st Annual IEEE Conference on Computational Complexity (CCC 2006), 16-20 July 2006, Prague, Czech Republic*, pages 88–95, 2006.

[BÖ0] Peter Bürgisser. *Completeness and Reduction in Algebraic Complexity Theory*, volume 7 of *Algorithms and Computation in Mathematics*. Springer-Verlag Berlin Heidelberg, 2000.

[BIJL18] Markus Bläser, Christian Ikenmeyer, Gorav Jindal, and Vladimir Lysikov. Generalized matrix completion and algebraic natural proofs. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018, Los Angeles, CA, USA, June 25-29, 2018*, pages 1193–1206, 2018.

[BIP16] Peter Bürgisser, Christian Ikenmeyer, and Greta Panova. No Occurrence Obstructions in Geometric Complexity Theory. In *IEEE 57th Annual Symposium on Foundations of Computer Science, FOCS 2016, 9-11 October 2016, Hyatt Regency, New Brunswick, New Jersey, USA*, pages 386–395, 2016.

[Bre76] R.P Brent. Multiple-precision zero-finding methods and the complexity of elementary function evaluation. *Analytic Computational Complexity*, pages 151–176, 1976.

[BS07] Andrej Bogdanov and Muli Safra. Hardness amplification for errorless heuristics. In *48th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2007), October 20-23, 2007, Providence, RI, USA, Proceedings*, pages 418–426, 2007.

[CKW11] Xi Chen, Neeraj Kayal, and Avi Wigderson. Partial derivatives in arithmetic complexity and beyond. *Foundations and Trends in Theoretical Computer Science*, 6(1-2):1–138, 2011.

[CLS18] Suryajith Chillara, Nutan Limaye, and Srikanth Srinivasan. Small-depth multilinear formula lower bounds for iterated matrix multiplication, with applications. In *35th Symposium on Theoretical Aspects of Computer Science, STACS 2018, February 28 to March 3, 2018, Caen, France*, pages 21:1–21:15, 2018.

[CM14] Suryajith Chillara and Partha Mukhopadhyay. Depth-4 lower bounds, determinantal complexity: A unified approach. In *31st International Symposium on Theoretical Aspects of Computer Science (STACS 2014), STACS 2014, March 5-8, 2014, Lyon, France*, pages 239–250, 2014.

[EGdOW18] Klim Efremenko, Ankit Garg, Rafael Mendes de Oliveira, and Avi Wigderson. Barriers for Rank Methods in Arithmetic Complexity. In *9th Innovations in Theoretical Computer Science Conference, ITCS 2018, January 11-14, 2018, Cambridge, MA, USA*, pages 1:1–1:19, 2018.

[FKS16] Michael A. Forbes, Mrinal Kumar, and Ramprasad Saptharishi. Functional lower bounds for arithmetic circuits and connections to boolean circuit complexity. In *31st Conference on Computational Complexity, CCC 2016, May 29 to June 1, 2016, Tokyo, Japan*, pages 33:1–33:19, 2016.

[FLMS15] Hervé Fournier, Nutan Limaye, Guillaume Malod, and Srikanth Srinivasan. Lower Bounds for Depth-4 Formulas Computing Iterated Matrix Multiplication. *SIAM J. Comput.*, 44(5):1173–1201, 2015.

[FPS08] Lance Fortnow, Aduri Pavan, and Samik Sengupta. Proving SAT does not have small circuits with an application to the two queries problem. *J. Comput. Syst. Sci.*, 74(3):358–363, 2008.

[Fro97] Georg Frobenius. Ueber die darstellung der endlichen gruppen durch linearc substitutionen. *Sitzungber. der Berliner Akademie*, 7:994–1015, 1897.

[FSS81] Merrick L. Furst, James B. Saxe, and Michael Sipser. Parity, circuits, and the polynomial-time hierarchy. In *22nd Annual Symposium on Foundations of Computer Science, Nashville, Tennessee, USA, 28-30 October 1981*, pages 260–270, 1981.

[FSV17] Michael A. Forbes, Amir Shpilka, and Ben Lee Volk. Succinct hitting sets and barriers to proving algebraic circuits lower bounds. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, Montreal, QC, Canada, June 19-23, 2017*, pages 653–664, 2017.

[Ges16] Fulvio Gesmundo. Gemetric aspects of iterated matrix multiplication. *Journal of Algebra*, 461:42–64, 2016.

[GGKS19] Ankit Garg, Nikhil Gupta, Neeraj Kayal, and Chandan Saha. Determinant equivalence test over finite fields and over Q. *Electronic Colloquium on Computational Complexity (ECCC)*, 26:42, 2019.

[GK98] Dima Grigoriev and Marek Karpinski. An exponential lower bound for depth 3 arithmetic circuits. In *Proceedings of the Thirtieth Annual ACM Symposium on the Theory of Computing, Dallas, Texas, USA, May 23-26, 1998*, pages 577–582, 1998.

[GKKS14] Ankit Gupta, Pritish Kamath, Neeraj Kayal, and Ramprasad Saptharishi. Approaching the Chasm at Depth Four. *J. ACM*, 61(6):33:1–33:16, 2014.

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

[Gro12] Joshua Abraham Grochow. *Symmetry and equivalence relations in classical and geometric complexity theory*. PhD thesis, Department of Computer Science, The University of Chicago, Chicago, Illinois, 2012.

[Hal15] Brian C Hall. *Lie Groups, Lie Algebras and Representations An Elementary introduction*. Springer, second edition, 2015.

[Hås86] Johan Håstad. Almost optimal lower bounds for small depth circuits. In *Proceedings of the 18th Annual ACM Symposium on Theory of Computing, May 28-30, 1986, Berkeley, California, USA*, pages 6–20, 1986.

[Hüt16] Jesko Hüttenhain. The Stabilizer of Elementary Symmetric Polynomials. *CoRR*, abs/1607.08419, 2016.

[IMW17] Christian Ikenmeyer, Ketan D. Mulmuley, and Michael Walter. On vanishing of kronecker coefficients. *Computational Complexity*, 26(4):949–992, 2017.

[IP16] Christian Ikenmeyer and Greta Panova. Rectangular kronecker coefficients and plethysms in geometric complexity theory. In *IEEE 57th Annual Symposium on Foundations of Computer Science, FOCS 2016, 9-11 October 2016, Hyatt Regency, New Brunswick, New Jersey, USA*, pages 396–405, 2016.

[Kay12] Neeraj Kayal. Affine projections of polynomials: extended abstract. In *Proceedings of the 44th Symposium on Theory of Computing Conference, STOC 2012, New York, NY, USA, May 19 - 22, 2012*, pages 643–662, 2012.

[KLSS17] Neeraj Kayal, Nutan Limaye, Chandan Saha, and Srikanth Srinivasan. An Exponential Lower Bound for Homogeneous Depth Four Arithmetic Formulas. *SIAM J. Comput.*, 46(1):307–335, 2017.

[KNST17] Neeraj Kayal, Vineet Nair, Chandan Saha, and Sébastien Tavenas. Reconstruction of Full Rank Algebraic Branching Programs. In *32nd Computational Complexity Conference, CCC 2017, July 6-9, 2017, Riga, Latvia*, pages 21:1–21:61, 2017.

[KS14a] Mrinal Kumar and Shubhangi Saraf. The limits of depth reduction for arithmetic formulas: it's all about the top fan-in. In *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*, pages 136–145, 2014.

[KS14b] Mrinal Kumar and Shubhangi Saraf. Superpolynomial lower bounds for general homogeneous depth 4 arithmetic circuits. In *Automata, Languages, and Programming - 41st International Colloquium, ICALP 2014, Copenhagen, Denmark, July 8-11, 2014, Proceedings, Part I*, pages 751–762, 2014.

[KS16a] Neeraj Kayal and Chandan Saha. Lower Bounds for Depth-Three Arithmetic Circuits with small bottom fanin. *Computational Complexity*, 25(2):419–454, 2016.

[KS16b] Mrinal Kumar and Shubhangi Saraf. Arithmetic circuits with locally low algebraic rank. In *31st Conference on Computational Complexity, CCC 2016, May 29 to June 1, 2016, Tokyo, Japan*, pages 34:1–34:27, 2016.

[KS16c] Mrinal Kumar and Shubhangi Saraf. Sums of products of polynomials in few variables: Lower bounds and polynomial identity testing. In *31st Conference on Computational Complexity, CCC 2016, May 29 to June 1, 2016, Tokyo, Japan*, pages 35:1–35:29, 2016.

[KS17a] Mrinal Kumar and Ramprasad Saptharishi. An exponential lower bound for homogeneous depth-5 circuits over finite fields. In *32nd Computational Complexity Conference, CCC 2017, July 6-9, 2017, Riga, Latvia*, pages 31:1–31:30, 2017.

[KS17b] Mrinal Kumar and Shubhangi Saraf. On the Power of Homogeneous Depth 4 Arithmetic Circuits. *SIAM J. Comput.*, 46(1):336–387, 2017.

[KSS14] Neeraj Kayal, Chandan Saha, and Ramprasad Saptharishi. A super-polynomial lower bound for regular arithmetic formulas. In *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*, pages 146–153, 2014.

[KST16a] Neeraj Kayal, Chandan Saha, and Sébastien Tavenas. An Almost Cubic Lower Bound for Depth Three Arithmetic Circuits. In *43rd International Colloquium on Automata, Languages, and Programming, ICALP 2016, July 11-15, 2016, Rome, Italy*, pages 33:1–33:15, 2016.

[KST16b] Neeraj Kayal, Chandan Saha, and Sébastien Tavenas. On the size of homogeneous and of depth four formulas with low individual degree. In *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2016, Cambridge, MA, USA, June 18-21, 2016*, pages 626–632, 2016.

[Lip89] Richard J. Lipton. New directions in testing. In *Distributed Computing And Cryptography, Proceedings of a DIMACS Workshop, Princeton, New Jersey, USA, October 4-6, 1989*, pages 191–202, 1989.

[MM62] Marvin Marcus and Francis May. The permanent function. *Canadian Journal of Mathematics*, 14:177–189, 1962.

[MR04] Thierry Mignon and Nicolas Ressayre. A quadratic bound for the determinant and permanent problem. *International Mathematics Research Notes*, 2004(79):4241–4253, 2004.

[MS01] Ketan Mulmuley and Milind A. Sohoni. Geometric complexity theory I: an approach to the P vs. NP and related problems. *SIAM J. Comput.*, 31(2):496–526, 2001.

[Mul99] Ketan Mulmuley. Lower bounds in a parallel model without bit operations. *SIAM J. Comput.*, 28(4):1460–1509, 1999.

[Mul07] Ketan Mulmuley. On P vs. NP, Geometric Complexity Theory, and the Flip I: a high level view. *CoRR*, abs/0709.0748, 2007.

[Mul10] Ketan Mulmuley. Explicit proofs and the flip. *CoRR*, abs/1009.0246, 2010.

[Mul11] Ketan Mulmuley. On P vs. NP and geometric complexity theory: Dedicated to Sri Ramakrishna. *J. ACM*, 58(2):5:1–5:26, 2011.

[Mul12] Ketan Mulmuley. The GCT program toward the *P* vs. *NP* problem. *Commun. ACM*, 55(6):98–107, 2012.

[MW18] Cody Murray and R. Ryan Williams. Circuit lower bounds for nondeterministic quasi-polytime: an easy witness lemma for NP and NQP. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018, Los Angeles, CA, USA, June 25-29, 2018*, pages 890–901, 2018.

[NW94] Noam Nisan and Avi Wigderson. Hardness vs randomness. *J. Comput. Syst. Sci.*, 49(2):149–167, 1994.

[NW97] Noam Nisan and Avi Wigderson. Lower Bounds on Arithmetic Circuits Via Partial Derivatives. *Computational Complexity*, 6(3):217–234, 1997.

[Raz85] Alexander A. Razborov. Lower bounds on the monotone complexity of some Boolean functions. *Soviet Mathematics Doklady*, 31:354–357, 1985.

[Raz87] Alexander A. Razborov. Lower bounds on the size of bounded-depth networks over a complete basis with logical addition. *Mathematical Notes of the Academy of Sciences of the USSR*, 41(4):333–338, 1987.

[Raz09] Ran Raz. Multi-linear formulas for permanent and determinant are of super-polynomial size. *J. ACM*, 56(2):8:1–8:17, 2009.

[Reg02] Kenneth W. Regan. Understanding the mulmuley-sohoni approach to P vs. NP. *Bulletin of the EATCS*, 78:86–99, 2002.

[RR97] Alexander A. Razborov and Steven Rudich. Natural Proofs. *J. Comput. Syst. Sci.*, 55(1):24–35, 1997.

[RY09] Ran Raz and Amir Yehudayoff. Lower Bounds and Separations for Constant Depth Multilinear Circuits. *Computational Complexity*, 18(2):171–207, 2009.

[Sch80] Jacob T. Schwartz. Fast Probabilistic Algorithms for Verification of Polynomial Identities. *J. ACM*, 27(4):701–717, 1980.

[Smo87] Roman Smolensky. Algebraic Methods in the Theory of Lower Bounds for Boolean Circuit Complexity. In *Proceedings of the 19th Annual ACM Symposium on Theory of Computing, 1987, New York, New York, USA*, pages 77–82, 1987.

[Val79] Leslie G. Valiant. Completeness classes in algebra. In *Proceedings of the 11h Annual ACM Symposium on Theory of Computing, April 30 - May 2, 1979, Atlanta, Georgia, USA*, pages 249–261, 1979.

[Wil14] Ryan Williams. Nonuniform ACC Circuit Lower Bounds. *J. ACM*, 61(1):2:1–2:32, 2014.

[Ye94] Yinyu Ye. Combining binary search and newton's method to compute real roots for a class of real functions. *J. Complexity*, 10(3):271–280, 1994.

[Zip79] Richard Zippel. Probabilistic algorithms for sparse polynomials. In *Symbolic and Algebraic Computation, EUROSAM '79, An International Symposiumon Symbolic and Algebraic Computation, Marseille, France, June 1979, Proceedings*, pages 216–226, 1979.

A Comparison between NW and the Permanent

The answers in Table 1 also hold over finite fields with mild restrictions on the characteristic and size of the field.

Questions	**Perm** (known results)	**NW** (Our Results)
1. Characterized by symmetries?	Yes over almost all fields	a. Yes over $\mathbb{C}$ b. No over $\mathbb{Q}, \mathbb{R}$.
2. Is an explicit basis of the Lie algebra known?	Yes	Yes
3. Is $\mathscr{G}_f$ generated by PS matrices?	Yes	Yes
4. Are all the diagonal symmetries continuous?	Yes over almost all fields	a. No over $\mathbb{C}$ b. Yes over $\mathbb{Q}, \mathbb{R}$.
5. Are all the permutation symmetries known?	Yes	Partially
6. Is a circuit testing algorithm known?	Yes	Yes
7. Is a Flip theorem known to hold?	Yes	Yes
8. Efficient equivalence test known?	Yes	Partially

Table 1: Comparison between NW and the Permanent

B Two observations on the design polynomial family

Observation B.1. *Let $k = k(d) \in [d]$ be an arbitrarily fixed,* poly(d)*-time computable, function of d. The design polynomial family $\mathcal{NW} := \{\mathsf{NW}_{d,k} : d \text{ is a prime}\}$ is in* VNP.

Proof. Owing to the density of primes, $\mathcal{NW}$ is a p-bounded family [Val79] as the number of variables and the degree of $\mathsf{NW}_{d,k}$ are both polynomial functions of d. By Proposition 2.20 of [BÖ0],

a p-bounded family $\{f_i\}_{i\in\mathbb{N}}$ is in VNP (i.e. p-definable) if the *coefficient computing function* for f_i is in #P. The coefficient computing function for f_i takes input a monomial in the variables of f_i and outputs the coefficient of the monomial in f_i. The coefficient computing function for $\mathsf{NW}_{d,k}$ can be shown to be in P as follows: Given a monomial m, check if it is set-multilinear in the sets $\mathbf{x}_0,\ldots,\mathbf{x}_{d-1}$. If not, the coefficient of m is 0 in $\mathsf{NW}_{d,k}$. Otherwise, let $m = x_{0,j_0}\cdots x_{d-1,j_{d-1}}$. Obtain a polynomial $h \in \mathbb{F}_d[z]_{d-1}$ by interpolating the points $(0,j_0),\ldots,(d-1,j_{d-1})$. Compute k from d. If $\deg(h) \le k$ then coefficient of m in $\mathsf{NW}_{d,k}$ is 1 else it is 0. □

Observation B.2. *Suppose $f \in \mathbb{F}[\mathbf{y}]$ is a degree-r polynomial having s monomials. Then, for $d \ge s$ and $d-k \ge r$, f is an affine projection of $\mathsf{NW}_{d,k}$.*

Proof. Fix a univariate $h \in \mathbb{F}_d[z]_k$ and set the variables $x_{0,h(0)},\ldots,x_{k-1,h(k-1)}$ to 1 and other variables of $\mathbf{x}_0,\ldots,\mathbf{x}_{k-1}$ to 0. The low-intersection property of $\mathsf{NW}_{d,k}$ ensures that under this setting, exactly d monomials remain in $\mathsf{NW}_{d,k}$. Moreover, these d monomials are pairwise variable disjoint and each monomial contains $d-k$ variables. As $d \ge s$ and $d-k \ge r$, we can map these d monomials to monomials of f via a simple substitution map from $\mathbf{x}$ to $\mathbf{y} \cup \mathbb{F}$. Hence, there is a $A \in \mathbb{F}^{n\times|\mathbf{y}|}$ and $\mathbf{b} \in \mathbb{F}^n$ such that $\mathsf{NW}_{d,k}(A\cdot\mathbf{y}+\mathbf{b}) = f$; in other words, f is an affine projection of $\mathsf{NW}_{d,k}$. □

C Proofs from Section 3

For a polynomial $p \in \mathbb{F}_d[z]$, m_p would refer to the monomial $\prod_{i\in[d]} x_{i,p(i)}$.

C.1 Proof of Claim 3.1

Claim 3.1 (restated): *The following matrices in $\mathbb{F}^{n\times n}$ are in $\mathscr{G}_{\mathsf{NW}}$:*

1. *A_β, a diagonal matrix with $((i,j),(i,j))$-th entry as $\beta_i \in \mathbb{F}^\times$ for $i,j \in [d]$, such that $\prod_{i\in[d]}\beta_i = 1$.*
2. *A_ℓ, a diagonal matrix with $((i,j),(i,j))$-th entry as $\zeta^{i^\ell\cdot j}$ for $i,j\in[d]$, where $\ell \in [d-k-1]$.*
3. *A_h, where $h \in \mathbb{F}_d[z]_k$, the $((i,j),(i,j+h(i)))$-th entry of A_h is 1 for $i,j\in[d]$ and other entries are 0.*

Proof. By definition, $A_\beta, A_\ell \in \mathrm{GL}_n(\mathbb{F})$. Also, $A_h \in \mathrm{GL}_n(F)$ as it is a permutation matrix. Observe that the polynomials $\mathsf{NW}(A_\beta\cdot\mathbf{x})$, $\mathsf{NW}(A_\ell\cdot\mathbf{x})$ and $\mathsf{NW}(A_h\cdot\mathbf{x})$ are obtained from $\mathsf{NW}(\mathbf{x})$ by replacing the variable $x_{i,j}$ with $\beta_i\cdot x_{i,j}$, $\zeta^{i^\ell\cdot j}\cdot x_{i,j}$ and $x_{i,j+h(i)}$ respectively, for $i,j\in[d]$. When A_β is applied on $\mathbf{x}$, a monomial m_p gets mapped to $\prod_{i\in[d]}\beta_i\cdot m_p = m_p$ as $\prod_{i\in[d]}\beta_i = 1$, implying $\mathsf{NW}(A_\beta\cdot\mathbf{x}) = \mathsf{NW}$. When A_h is applied on $\mathbf{x}$, a monomial m_p gets mapped to m_{p+h}; in other words, the monomials of NW are 'shifted around' and so $\mathsf{NW}(A_h\cdot\mathbf{x}) = \mathsf{NW}$. When A_ℓ is applied on $\mathbf{x}$, a monomial m_p is mapped to $\prod_{i\in[d]}\zeta^{i^\ell\cdot p(i)}\cdot m_p$. We show below that $\prod_{i\in[d]}\zeta^{i^\ell\cdot p(i)} = 1$ for every $\ell\in[d-k-1]$, thereby implying $\mathsf{NW}(A_\ell\cdot\mathbf{x}) = \mathsf{NW}$.

Observation C.1. *For every $p\in\mathbb{F}_d[\mathbf{x}]_k$ and $\ell\in[d-k-1]$, $\prod_{i\in[d]}\zeta^{i^\ell\cdot p(i)} = 1$.*

Proof. As $\zeta\neq 1$ is a d-th root of unity, $\prod_{i\in[d]}\zeta^{i^\ell\cdot p(i)} = \zeta^{\sum_{i\in\mathbb{F}_d} i^\ell\cdot p(i)}$ and so it is sufficient to show that $\sum_{i\in\mathbb{F}_d} i^\ell\cdot p(i) = 0$. Suppose $p(z) = a_r z^r+\cdots+a_0$, where $r\le k$ and $a_r,\ldots,a_0\in\mathbb{F}_d$. Then

$$\sum_{i\in\mathbb{F}_d} i^\ell\cdot p(i) = a_r\left(\sum_{i\in\mathbb{F}_d} i^{r+\ell}\right)+\cdots+a_0\left(\sum_{i\in\mathbb{F}_d} i^\ell\right).$$

Each summand in the RHS of the above equation is of the form $a \cdot \left(\sum_{i \in \mathbb{F}_d} i^s\right)$, where $0 \leq s \leq d-2$. As $\sum_{i \in \mathbb{F}_d} i^0 = 0$, assume that $1 \leq s \leq d-2$. Let b be a generator of $\mathbb{F}_d^\times$. Then

$$\sum_{i \in \mathbb{F}_d} i^s = \sum_{i \in \mathbb{F}_d^\times} i^s = \sum_{t \in [d-1]} b^{t \cdot s} = \frac{1 - b^{(d-1)\cdot s}}{1 - b^s} = 0, \quad \text{as } b^{d-1} = 1 \text{ in } \mathbb{F}_d. \tag{2}$$

Hence, $\sum_{i \in \mathbb{F}_d} i^\ell \cdot p(i) = 0$ implying $\prod_{i \in [d]} \zeta^{i^\ell \cdot p(i)} = 1$. □

Thus, A_β, A_ℓ and A_h belong to $\mathscr{G}_{\mathsf{NW}}$ over $\mathbb{F}$. □

C.2 Proof of Claim 3.2

Claim 3.2 (restated): *Let f be a homogeneous degree-d polynomial in $\mathbb{F}[\mathbf{x}]$. If $\mathscr{G}_f$ contains the matrices A_β, A_ℓ and A_h (for all choices of β, ℓ and h, as mentioned in Claim 3.1) then $f = \alpha \cdot \mathsf{NW}$ for some $\alpha \in \mathbb{F}$.*

Proof. Let $f \neq 0$, otherwise we have nothing to prove. The presence of A_β in $\mathscr{G}_f$ implies that f is a set-multilinear polynomial with respect to the partition $\biguplus_{i \in [d]} \mathbf{x}_i$. If not then there is a term $\alpha \cdot m$ in f, where $\alpha \in \mathbb{F}^\times$ and m is a degree-d monomial with no $\mathbf{x}_t$-variables for some $t \in [d]$. Pick a $\gamma \in \mathbb{F}^\times$ such that $\gamma^d \neq 1$ [13]. Now, set $\beta_i = \gamma$ for $i \in [d] \backslash \{t\}$ and $\beta_t = \gamma^{-(d-1)}$ so that $\prod_{i \in [d]} \beta_i = 1$ is satisfied. When A_β is applied on $\mathbf{x}$, the term $\alpha \cdot m$ maps to $\alpha\gamma^d \cdot m \neq \alpha \cdot m$, implying that $f(A_\beta \cdot \mathbf{x}) \neq f(\mathbf{x})$.

As f is set-multilinear, every term of f is of the kind $\alpha_p \cdot m_p$, where $\alpha_p \in \mathbb{F}^\times$ and $p \in \mathbb{F}_d[z]$ with $\deg(p) \leq d-1$. This is because any function from $\mathbb{F}_d$ to $\mathbb{F}_d$ can be represented by a univariate polynomial of degree at most $d-1$. We now show that $\deg(p) \leq k$ for every term $\alpha_p \cdot m_p$ in f. Suppose not. Then, there is a term $\alpha_p \cdot m_p$ such that $p = a_r z^r + \cdots + a_0$, $r > k$ and $a_r \neq 0$. When A_ℓ is applied on $\mathbf{x}$, the term $\alpha_p \cdot m_p$ gets mapped to $\prod_{i \in [d]} \zeta^{i^\ell \cdot p(i)} \cdot \alpha_p \cdot m_p$. Now choose $\ell = d - r - 1 \leq d - k - 2$. That $\prod_{i \in [d]} \zeta^{i^\ell \cdot p(i)} \neq 1$ for this choice of ℓ can be argued as follows: Since $\prod_{i \in [d]} \zeta^{i^\ell \cdot p(i)} = \zeta^{\sum_{i \in \mathbb{F}_d} i^\ell \cdot p(i)}$, it is sufficient to show that $\sum_{i \in \mathbb{F}_d} i^\ell \cdot p(i) \neq 0$. Expanding the sum,

$$\sum_{i \in \mathbb{F}_d} i^\ell \cdot p(i) = a_r \left(\sum_{i \in \mathbb{F}_d} i^{d-1}\right) + a_{r-1} \left(\sum_{i \in \mathbb{F}_d} i^{d-2}\right) + \cdots + a_0 \left(\sum_{i \in \mathbb{F}_d} i^{d-r-1}\right).$$

As argued in Equation (2), the above sum is $a_r \cdot (d-1) \neq 0$, implying $f(A_\ell \cdot \mathbf{x}) \neq f(\mathbf{x})$. Hence, every term $\alpha_p \cdot m_p$ of f must have $\deg(p) \leq k$.

When A_h is applied on $\mathbf{x}$, a term $\alpha_p \cdot m_p$ maps to $\alpha_p \cdot m_{p+h}$ which implies $\alpha_p = \alpha_{p+h}$. Running over all $h \in \mathbb{F}_d[z]_k$, we get $\alpha_p = \alpha$ for every $p \in \mathbb{F}_d[z]_k$, for some $\alpha \in \mathbb{F}^\times$. Hence, $f = \alpha \cdot \mathsf{NW}$.

□

C.3 Proof of Lemma 3.1

Lemma 3.1 (restated): *Polynomial* NW *is characterized by circuit identities over any field $\mathbb{F}$.*

[13] As $|\mathbb{F}| \neq d+1$, such a γ always exists.

Proof. Recall, $n = |\mathbf{x}| = d^2$. We show that if an n-variate polynomial $f \in \mathbb{F}[\mathbf{x}]$ satisfies the following polynomial identities then $f = \alpha \cdot \mathsf{NW}$ for some $\alpha \in \mathbb{F}$. The rows and columns of the $n \times n$ matrices in the identities below are indexed by the set $\{(i,j) : i,j \in \mathbb{F}_d\}$.

1. $q_1(f(A_i(u) \cdot \mathbf{x}), f(\mathbf{x}), u) = 0$, for $i \in [d]$, where $q_1(z_1, z_2, u) := z_1 - u \cdot z_2$. Here, $A_i(u) \in \mathbb{F}[u]^{n \times n}$ is a diagonal matrix with the $((i,j),(i,j))$-th entry as u, for every $j \in [d]$, and the other diagonal entries as 1.

2. $q_2(f(A_{a,r} \cdot \mathbf{x}), f(\mathbf{x})) = 0$, for $a \in \mathbb{F}_d^\times$ and $r \in [k+1]$, where $q_2(z_1, z_2) := z_1 - z_2$. Here, $A_{a,r} \in \mathbb{F}^{n \times n}$ with the $((i,j),(i,j+a \cdot i^r))$-th entry as 1, for every $i,j \in \mathbb{F}_d$, and the other entries as 0.

3. $q_3(f(A_t \cdot \mathbf{x})) = 0$, for $t \in [d] \backslash [k+1]$, where $q_3(z) := z$. Here, $A_t \in \mathbb{F}^{n \times n}$ is a diagonal matrix with the $((t,0),(t,0))$-th and the $((i,j),(i,j))$-th entries as 0, for every $i \in [k+1], j \in [d] \backslash \{0\}$, and the remaining diagonal entries as 1.

Observe that there are $\mathrm{poly}(n)$ many identities above: d many under item 1, $(d-1)(k+1)$ many under item 2, and $(d-k-1)$ many under item 3. Also, it is clear that every q_i is computable by a constant size circuit, and the matrices $A_i(u), A_{a,r}$ and A_t are computable by $\mathrm{poly}(n)$ size circuits. The identities under item 1 imply that f is a set-multilinear, homogeneous, degree-d polynomial. If not then f contains a term $\beta \cdot m$, where the degree of the $\mathbf{x}_i$-variables in m is $e \neq 1$ for some $i \in [d]$. On applying $A_i(u)$ to $\mathbf{x}$, the term $\beta \cdot m$ gets mapped to $u^e \beta \cdot m \neq u\beta \cdot m$, implying $f(A_i(u) \cdot \mathbf{x}) \neq u \cdot f(\mathbf{x})$, i.e., $q_1(f(A_i(u) \cdot \mathbf{x}), f(\mathbf{x}), u) \neq 0$.

As f is set-multilinear and homogeneous, every term of f looks like $\alpha_p \cdot m_p$, where $\alpha_p \in \mathbb{F}^\times$ and $m_p = \prod_{i \in \mathbb{F}_d} x_{i,p(i)}$ for some $p \in \mathbb{F}_d[z]$ with $\deg(p) \leq d-1$. When $A_{a,r}$ is applied on $\mathbf{x}$, for some $a \in \mathbb{F}_d^\times$ and $r \in [k+1]$, a term $\alpha_p \cdot m_p$ maps to $\alpha_p \cdot m_{p+h}$, where $h = az^r \in \mathbb{F}_d[z]_k$. Since, f satisfies the identities in item 2, $f(A_{a,r} \cdot \mathbf{x}) = f(\mathbf{x})$ and so $\alpha_p \cdot m_{p+h}$ is also a term in f. By varying $a \in \mathbb{F}_d^\times$ and $r \in [k+1]$, we see that f contains the term $\alpha_p \cdot m_{p+h}$ for every $h \in \mathbb{F}_d[z]_k$. Thus, there is a set $\mathcal{S} \subseteq \mathbb{F}_d[z]_{d-1}$ such that f is of the form,

$$f = \sum_{p \in \mathcal{S}} \alpha_p \cdot \sum_{h \in \mathbb{F}_d[z]_k} m_{p+h}. \tag{3}$$

If $f \neq \alpha \cdot \mathsf{NW}$ for all $\alpha \in \mathbb{F}$, then there is a $p \in \mathbb{F}_d[z]$ with $\deg(p) > k$ such that f contains a term $\alpha_p \cdot m_p$ for some $\alpha_p \in \mathbb{F}^\times$. Let h be the polynomial in $\mathbb{F}_d[z]_k$ such that $h(i) = -p(i)$ for all $i \in [k+1]$. From Equation (3), f contains the term $\alpha_p \cdot m_{p+h}$. As $\deg(p) > k$, $h(z) \neq -p(z)$. So, there is a $t \in [d] \backslash [k+1]$ such that $p(t) + h(t) \neq 0$. On applying A_t to $\mathbf{x}$, only those terms of f survive that contain the variables $x_{0,0}, \ldots, x_{k,0}$ but do not contain $x_{t,0}$, and $\alpha_p \cdot m_{p+h}$ is such a term. Hence, $q_3(f(A_t \cdot \mathbf{x})) = f(A_t \cdot \mathbf{x}) \neq 0$. This contradicts f satisfying the identities in item 3. Therefore, $f = \alpha \cdot \mathsf{NW}$, for some $\alpha \in \mathbb{F}$. On the other hand, any $f = \alpha \cdot \mathsf{NW}$ satisfies all the identities. □

D Proofs from Section 4

D.1 Proof of Lemma 4.1

Lemma 4.1 (restated): *Let $\mathbb{F}$ be a field and* $\mathrm{char}(\mathbb{F}) \neq d$. *The dimension of $\mathfrak{g}_{\mathsf{NW}}$ over $\mathbb{F}$ is $d-1$, and the diagonal matrices $B_1, \ldots, B_\ell$ (defined below) form a $\mathbb{F}$-basis of $\mathfrak{g}_{\mathsf{NW}}$. For $\ell \in \{1, \ldots, d-1\}$,*

$$(B_\ell)_{(i,j),(i,j)} = \begin{cases} 1, & \text{if } i = 0, j \in [d] \\ -1, & \text{if } i = \ell, j \in [d] \\ 0, & \text{otherwise.} \end{cases}$$

Proof. Recall that the rows and columns of a matrix in $\mathfrak{g}_{\mathsf{NW}}$ are indexed by the set $\{(i,j) : i, j \in \mathbb{F}_d\}$. By Definition 2.7, $B = (\alpha_{(i,j),(l,r)})_{i,j,l,r \in [d]} \in \mathfrak{g}_{\mathsf{NW}}$ if and only if the following equation is satisfied:

$$\sum_{i,j,l,r \in [d]} \alpha_{(i,j),(l,r)} \cdot x_{l,r} \cdot \partial_{ij}\mathsf{NW} = 0, \qquad \text{where } \partial_{ij}\mathsf{NW} := \frac{\partial \mathsf{NW}}{\partial x_{i,j}}. \tag{4}$$

Claim D.1. *Every $B = (\alpha_{(i,j),(l,r)})_{i,j,l,r \in [d]} \in \mathfrak{g}_{\mathsf{NW}}$ is a diagonal matrix.*

Proof. Let $i, j, l, r \in [d]$, such that $(i,j) \neq (l,r)$. It follows from the low-intersection property of NW that the terms $x_{l,r} \cdot \partial_{ij}\mathsf{NW}$ and $x_{u,v} \cdot \partial_{st}\mathsf{NW}$ in Equation (4) are monomial disjoint for every $s, t, u, v \in [d]$ satisfying $(s,t) \neq (i,j)$ or $(u,v) \neq (l,r)$. Hence, $\alpha_{(i,j),(l,r)} = 0$ and B is diagonal. □

Thus, $\mathfrak{g}_{\mathsf{NW}}$ can be viewed as a subspace of $\mathbb{F}^n$ by associating a column vector $\mathbf{w}_B := B \cdot \mathbf{1} \in \mathbb{F}^n$ with every $B \in \mathfrak{g}_{\mathsf{NW}}$, where $\mathbf{1}$ is the all-one column vector in $\mathbb{F}^n$. The coordinates of $\mathbf{w}_B$ are indexed by $\{(l,r) : l, r \in \mathbb{F}_d\}$ and $\mathbf{w}_B(l,r) = \alpha_{(l,r),(l,r)}$ is its (l,r)-th coordinate. Now, we construct a matrix $D \in \mathbb{F}^{n \times n}$ using degree-0 and degree-1 polynomials in $\mathbb{F}_d[z]_k$, such that $\mathfrak{g}_{\mathsf{NW}}$ (viewed as a subspace of $\mathbb{F}^n$) is contained in $\mathrm{Ker}_{\mathbb{F}}(D)$, the kernel of D [14]. This would help us find $\dim_{\mathbb{F}}(\mathfrak{g}_{\mathsf{NW}})$.

Construction of matrix D. The rows of D are indexed by $\{(a,b) : a, b \in \mathbb{F}_d\}$, where (a,b) corresponds to the univariate $bz + a \in \mathbb{F}_d[z]$. The columns are indexed by $\{(l,r) : l, r \in \mathbb{F}_d\}$, where (l,r) corresponds to the variable $x_{l,r}$ (as before). D is a $0/1$ matrix. The $((a,b),(l,r))$-th entry of D is 1 if $x_{l,r}$ is present in the monomial $\prod_{i \in \mathbb{F}_d} x_{i,bi+a}$, else it is 0. Denote the (a,b)-th row of D by R_{ab}. We record a few easy-to-verify properties of D below:

1. For $a, b \in [d]$, R_{ab} contains d many 1.
2. For every $a \in [d]$, the rows $\{\mathsf{R}_{ab} : b \in [d]\}$ contain 1 in the $(0,a)$-th column and 0 in the columns indexed by $(0,r)$ where $r \neq a$.
3. Let $B \in \mathfrak{g}_{\mathsf{NW}}$ and $\mathbf{w}_B(l,r) = \alpha_{(l,r),(l,r)}$ for $l, r \in \mathbb{F}_d$. Then, $(D \cdot \mathbf{w}_B)(a,b) = \sum_{l \in [d]} \alpha_{(l,bl+a),(l,bl+a)}$, which is the coefficient of monomial $\prod_{l \in \mathbb{F}_d} x_{l,bl+a}$ in the LHS of Equation (4). This implies $(D \cdot \mathbf{w}_B)(a,b) = 0$ for every $a, b \in [d]$, and hence, $\mathbf{w}_B \in \mathrm{Ker}_{\mathbb{F}}(D)$.

[14] Matrix D would just be a part of the coefficient matrix of the linear system obtained from the equations $\sum_{l \in [d]} \alpha_{(l,h(l)),(l,h(l))} = 0$, for all $h \in \mathbb{F}_d[z]_k$. Here, $\{\alpha_{(l,r),(l,r)} : l, r \in \mathbb{F}_d\}$ are the d^2 variables of the system.

We argue that the rank of D is at least d^2-d+1, by showing the $\mathbb{F}$-linear independence of the rows indexed by $\{(a,b) : a \in [d-1], b \in [d]\}$ and $(d-1,0)$. This, along with property 3, would imply that $\dim_{\mathbb{F}}(\mathfrak{g}_{\mathsf{NW}}) \leq d-1$. Property 2 implies that it is sufficient to show the $\mathbb{F}$-linear independence of the d^2-d rows indexed by $\{(a,b) : a \in [d-1], b \in [d]\}$, as the row indexed by $(d-1,0)$ contains 1 in the column indexed by $(0,d-1)$ and this column contains 0 in the rows indexed by $\{(a,b) : a \in [d-1], b \in [d]\}$.

Claim D.2. *The rows $\{\mathsf{R}_{ab} : a \in [d-1], b \in [d]\}$ are $\mathbb{F}$-linearly independent, if $\mathrm{char}(\mathbb{F}) \neq d$.*

Proof. We multiply these rows with formal variables $\Gamma := \{\gamma_{ab} : a \in [d-1], b \in [d]\}$, and show that if the following equation holds then each $\gamma_{ab} = 0$. The number of Γ-variables is $|\Gamma| = d^2-d$.

$$\sum_{a\in[d-1],b\in[d]} \gamma_{ab} \cdot \mathsf{R}_{ab} = 0.$$

From the above equation, we get d^2 linear equations in the Γ-variables, one for every coordinate of the rows. Fix $a \in [d-1]$ and $b \in [d]$ arbitrarily. From property 1, there are exactly d equations (one for each $l \in [d]$) containing the variable γ_{ab}. We can naturally identify these d equations with $l \in [d]$. The variables γ_{ab} and $\gamma_{a'b'}$ are present in the equation corresponding to a $l \in [d]$ if and only if $bl + a = b'l + a'$ over $\mathbb{F}_d$. Equation (5) corresponds to $l = 0$ and Equation (6) corresponds to a $l \in [d]\backslash\{0\}$.

$$\gamma_{a\,0} + \cdots + \gamma_{a\,b} + \cdots + \gamma_{a\,d-1} = 0, \tag{5}$$

$$\left(\sum_{a'\in[d-1]\setminus\{a\}} \gamma_{a'b'}\right) + \gamma_{ab} = 0, \qquad \text{where } b' = b + \frac{a-a'}{l}. \tag{6}$$

For $a \in [d-1]$ and $b, l \in [d]$, denote the linear forms at the LHS of these equations as 'Equation $(5)_{a,b}$' and 'Equation $(6)_{a,b,l}$'. A simple counting argument imply the following.

Observation D.1. *Let $a \in [d-1]$ and $b \in [d]$. Consider the d linear forms, Equation $(5)_{a,b}$ and Equation $(6)_{a,b,l}$ for $l \in [d]\backslash\{0\}$. Every pair of these d linear forms has γ_{ab} as the only common Γ-variable. Further, these d linear forms together contain all the Γ-variables except the variables in $\{\gamma_{a'b} : a' \in [d-1] \setminus \{a\}\}$.*

The next two observations will help us conclude that $\gamma_{ab} = 0$.

Observation D.2. *Let $a \in [d-1]$, $b \in [d]$ and $b' \in [d] \setminus \{b\}$. There is exactly one linear form in $\{$Equation $(6)_{a,b',l} : l \in [d] \setminus \{0\}\}$ that contains no Γ-variable from $\{\gamma_{a'b} : a' \in [d-1] \setminus \{a\}\}$. This unique linear form is Equation $(6)_{a,b',l(b')}$, where $l(b') = \frac{(d-1)-a}{b'-b}$.*

Proof. The linear form Equation $(6)_{a,b',l}$ contains a variable $\gamma_{a'b}$ if and only if $l = \frac{a'-a}{b'-b}$. If we choose $l = l(b') = \frac{(d-1)-a}{b'-b}$ then a' is forced to take value $d-1$. Thus, Equation $(6)_{a,b',l(b')}$ contains no variable from $\{\gamma_{a'b} : a' \in [d-1] \setminus \{a\}\}$. On the other hand, for $l \in [d]\backslash\{0\}$ and $l \neq l(b')$, there is exactly one variable in $\{\gamma_{a'b} : a' \in [d-1] \setminus \{a\}\}$ that belongs to Equation $(6)_{a,b',l}$. □

With $l(b')$ defined as above, we have the following observation.

Observation D.3. *Let $a \in [d-1]$, $b \in [d]$ and b', b'' be two distinct elements in $[d] \setminus \{b\}$. The linear forms Equation $(6)_{a,b',l(b')}$ and Equation $(6)_{a,b'',l(b'')}$ do not have any Γ-variable in common.*

Proof. For contradiction, suppose $\gamma_{\tilde{a}\tilde{b}}$ appears in both Equation $(6)_{a,b',l(b')}$ and Equation $(6)_{a,b'',l(b'')}$. Then, $\tilde{b} = b' + \frac{a-\tilde{a}}{l(b')} = b'' + \frac{a-\tilde{a}}{l(b'')}$. Hence,

$$b' - b'' = (a - \tilde{a}) \cdot \left(\frac{1}{l(b'')} - \frac{1}{l(b')} \right) = (a - \tilde{a}) \cdot \frac{b'' - b'}{(d-1) - a} \quad ,$$

by plugging in the values of $l(b')$ and $l(b'')$. As $\tilde{a} \neq d-1$, the above equality cannot hold. □

Finally, consider the following equation, which is implied from Equations (5) and (6),

$$\text{Equation } (5)_{a,b} \; + \sum_{l \in [d] \setminus \{0\}} \text{Equation } (6)_{a,b,l} \; - \sum_{b' \in [d] \setminus \{b\}} \text{Equation } (6)_{a,b',l(b')} \; = \; 0.$$

By Observation D.1, Equation $(5)_{a,b} \; + \; \sum_{l \in [d] \setminus \{0\}}$ Equation $(6)_{a,b,l}$ is the sum of $d \cdot \gamma_{ab}$ and all the Γ-variables barring $\{\gamma_{a'b} : a' \in [d-1] \setminus \{a\}\} \uplus \{\gamma_{ab}\}$. On the other hand, Observations D.2 and D.3 and a simple counting argument, imply that $\sum_{b' \in [d] \setminus \{b\}}$ Equation $(6)_{a,b',l(b')}$ is the sum of all the Γ-variables barring $\{\gamma_{a'b} : a' \in [d-1] \setminus \{a\}\} \uplus \{\gamma_{ab}\}$. Therefore, $\gamma_{ab} = 0$ as $\text{char}(\mathbb{F}_d) \neq d$. This proves the $\mathbb{F}$-linear independence of $\{\mathsf{R}_{ab} : a \in [d-1], b \in [d]\}$. □

Thus, we have shown that $\dim_{\mathbb{F}}(\mathfrak{g}_{\mathsf{NW}}) \leq d-1$. This immediately implies that $\dim_{\mathbb{F}}(\mathfrak{g}_{\mathsf{NW}}) = d-1$, as the matrices $B_1, \ldots, B_{d-1}$ (in the statement of Lemma 4.1) are $\mathbb{F}$-linearly independent and they belong to $\mathfrak{g}_{\mathsf{NW}}$ (as they satisfy Equation (4)). Lemma 4.1 implies the following corollary. We would see a version of this corollary over a commutative ring $\mathcal{R}$ in Claim D.4.

Corollary D.1. *Let $\mathbb{F}$ be a field such that $\text{char}(\mathbb{F}) \neq d$. Consider the linear system over $\mathbb{F}$ obtained from the equations $\sum_{i \in [d]} x_{i,h(i)} = 0$ for all $h \in \mathbb{F}_d[z]_k$, where $\{x_{i,j} : i,j \in [d]\}$ are the variables. The solution space of the system consists of the solutions $x_{i,0} = x_{i,1} = \ldots = x_{i,d-1} = \alpha_i$ for every $i \in [d]$, where $\alpha_0, \ldots, \alpha_{d-1} \in \mathbb{F}$ satisfy $\sum_{i \in [d]} \alpha_i = 0$, and these are the only solutions.*

□

D.1.1 More on matrix D

Let $\mathbb{F}$ be a field such that $\text{char}(\mathbb{F}) \neq d$. Suppose $u = d^2 - d + 1$ and $\mathbf{x}' = \mathbf{x} \setminus \{x_{1,0}, \ldots, x_{d-1,0}\}$. We know from Lemma 4.1 that the first u rows of D are $\mathbb{F}$-linearly independent. Let B be the $u \times d^2$ size matrix obtained by restricting D to the first u rows, which are indexed by the polynomials $\{az + b \; : \; a \in [d], b \in [d-1]\} \cup \{d-1\}$. Further, let C be the $u \times u$ matrix obtained by restricting B to the columns indexed by $\mathbf{x}'$.

Claim D.3. *The absolute value of the determinant of C over $\mathbb{Z}$ is d^r, where $r = O(d^2)$.*

Proof. We know that

$$B \cdot \mathbf{x} = \Big(\sum_{i \in \mathbb{F}_d} x_{i,0} \cdots \sum_{i \in \mathbb{F}_d} x_{i,(d-1)i} \cdots \sum_{i \in \mathbb{F}_d} x_{i,d-2} \cdots \sum_{i \in \mathbb{F}_d} x_{i,(d-1)i+(d-2)} \sum_{i \in \mathbb{F}_d} x_{i,(d-1)} \Big)^T,$$

which gives the following set of linear polynomials in $\mathbf{x}$.

$$S_1 = \left\{ \sum_{i \in \mathbb{F}_d} x_{i,h(i)} \; : \; h \in \{az + b \; : \; a \in [d], b \in [d-1]\} \cup \{d-1\} \right\}.$$

Let S_2 be the set of $d^2 - d + 1$ distinct linear polynomials in $\mathbf{x}$ defined as

$$S_2 = \{x_{i,j} - x_{i,0} : i \in [d], j \in [d] \setminus \{0\}\} \cup \left\{ \sum_{i \in \mathbb{F}_d} x_{i,0} \right\}.$$

Consider the following fact.

Fact 1. *Suppose S_1, S_2 are two sets of linear polynomials in n variables over $\mathbb{F}$ having same solution spaces. Then,* $\text{span}_{\mathbb{F}}\{S_1\} = \text{span}_{\mathbb{F}}\{S_2\}$.

Then, from Corollary D.1 and Fact 1, we get

$$\text{span}_{\mathbb{F}}\{S_1\} = \text{span}_{\mathbb{F}}\{S_2\}. \tag{7}$$

Let A be the $u \times d^2$ coefficient matrix of the polynomials in S_2. Then, Equation (7) implies that there is a $M \in \text{GL}_u(\mathbb{F})$ such that

$$M \cdot B = A. \tag{8}$$

Let A_1 be the $u \times u$ matrix obtained by restricting A to the columns indexed by $\mathbf{x}'$. It is easy to see that A_1 is invertible. Hence, Equation (8) implies that $M \cdot C = A_1$ and so C is also invertible.

We claim that $\det_{\mathbb{Z}}(C) = d^r$ for some $r \in \mathbb{N}$. Suppose not. Then there exists a prime number $p \neq d$ such that $p \mid \det_{\mathbb{Z}}(C)$. Then, the determinant of C is 0 over the finite field $\mathbb{F}_p$, which is a contradiction as $\text{char}(\mathbb{F}_p) \neq d$. As C is a 0/1 matrix, $|\det_{\mathbb{Z}}(C)| \leq (d^2 - d + 1)!$, which implies $r = O(d^2)$. □

Claim D.3 implies the following.

Claim D.4. *Let $\mathcal{R}$ be a ring with multiplicative identity such that d is invertible in $\mathcal{R}$. Consider the linear system over $\mathcal{R}$ obtained from the equations $\sum_{i \in [d]} x_{i,h(i)} = 0$ for all $h \in \mathbb{F}_d[z]_k$, where $\{x_{i,j} : i, j \in [d]\}$ are the variables. The solution space of the system consists of the solutions $x_{i,0} = x_{i,1} = \ldots = x_{i,d-1} = \alpha_i$ for every $i \in [d]$, where $\alpha_0, \ldots, \alpha_{d-1} \in \mathcal{R}$ satisfy $\sum_{i \in [d]} \alpha_i = 0$, and these are the only solutions.*

Proof. Recall the definitions of matrices B and C given in the beginning of this section. Observe that $B \cdot \mathbf{x} = 0$ implies the following

$$C \cdot \mathbf{x}' = \mathbf{v},$$

where the entries of $\mathbf{v}$ are linear forms in $x_{1,0}, \ldots, x_{d-1,0}$. Let $\text{Adj}(C)$ be the adjoint of C. (Observe that entries of $\text{Adj}(C)$ are integers and are well-defined in $\mathcal{R}$.) On multiplying the above equation with $\text{Adj}(C)$, we get

$$\text{Adj}(C) \cdot C \cdot \mathbf{x}' = \text{Adj}(C) \cdot \mathbf{v},$$

which implies

$$\det(C) \cdot \mathbf{x}' = \mathbf{v}', \tag{9}$$

where $\mathbf{v}' = \text{Adj}(C) \cdot \mathbf{v}$. Clearly, every entry of $\mathbf{v}'$ is a linear form in $x_{1,0}, \ldots, x_{d-1,0}$. This equation holds over any commutative ring $\mathcal{R}$ with multiplicative identity. In particular, it also holds over a field $\mathbb{F}$ such that $\text{char}(\mathbb{F}) \neq d$. From Corollary D.1, we know $x_{i,0} = x_{i,1} = \ldots = x_{i,d-1}$ for every $i \in [d]$ and $\sum_{i \in [d]} x_{i,0} = 0$. Thus, for $i \in \{1, \ldots, d-1\}, j \in [d] \setminus \{0\}$ the entry of $\mathbf{v}'$ indexed by $x_{i,j}$ must be $\det(C) \cdot x_{i,0}$, and for $j \in [d]$ the entry indexed by $x_{0,j}$ must be $\det(C) \cdot (-(\sum_{i=1}^{d-1} x_{i,0}))$. From Claim D.3, we know that $\det(C) = d^r$. As d is invertible in $\mathcal{R}$, on multiplying Equation (9) with $(\det(C))^{-1}$, we get the result. □

D.1.2 A corollary of Lemma 4.1

Corollary D.2. *If* $|\mathbb{F}| > \binom{d}{2}$ *then there exists a* $B = \mathrm{diag}(\alpha_0, \ldots, \alpha_d) \otimes I_d \in \mathfrak{g}_{\mathsf{NW}}$ *such that* $\alpha_0, \ldots, \alpha_{d-1}$ *are distinct elements of* $\mathbb{F}$ *and* $\sum_{i\in[d]} \alpha_i = 0$.

Proof. Treat $\alpha_0, \ldots, \alpha_{d-2}$ as formal variables and let $\alpha_{d-1} = -(\alpha_0 + \cdots + \alpha_{d-2})$. By the Schwartz-Zippel lemma,

$$\Pr_{\alpha_0,\ldots,\alpha_{d-2}\in_r\mathbb{F}}[\text{there exist } i,l \in [d] \text{ such that } i \neq l \text{ and } \alpha_i = \alpha_l] \leq \frac{\binom{d}{2}}{|\mathbb{F}|} < 1.$$

Hence, there exists such a $B \in \mathfrak{g}_{\mathsf{NW}}$. □

D.2 Proof of Claim 4.1

Let $A \in \mathscr{G}_{\mathsf{NW}}$. For $i, l \in [d]$, the (i,l)-th *block of* A, denoted A_{il}, is a sub-matrix of A whose rows are indexed by the set $\{(i,j) : j \in [d]\}$ (called the i-th block of rows) and columns indexed by $\{(l,j) : j \in [d]\}$ (called the l-th block of columns).

Claim 4.1 (restated): *Every* $A \in \mathscr{G}_{\mathsf{NW}}$ *is a block-permuted matrix with block size* d.

Proof. Choose a $B \in \mathfrak{g}_{\mathsf{NW}}$ arbitrarily. From Lemma 2.2, there exists a $C \in \mathfrak{g}_{\mathsf{NW}}$ such that

$$A \cdot C = B \cdot A.$$

From Lemma 4.1, $B = \mathrm{diag}(\alpha_0, \ldots, \alpha_{d-1}) \otimes I_d$ and $C = \mathrm{diag}(\gamma_0, \ldots, \gamma_{d-1}) \otimes I_d$, where $\sum_{i\in[d]} \alpha_i = \sum_{i\in[d]} \gamma_i = 0$. The above equation implies, for every $i, l \in [d]$,

$$\gamma_l \cdot A_{il} = \alpha_i \cdot A_{il},$$

where A_{il} is the (i,l)-th block of A. If A is not block-permuted then for some $l \in [d]$, there are non-zero blocks A_{il} and $A_{i'l}$ such that $i \neq i'$ (as A is non-singular). For this choice of l, i and i', the last equation implies $\gamma_l = \alpha_i = \alpha_{i'}$. This contradicts Corollary D.2, as B is chosen arbitrarily. □

D.3 Proof of Theorem 5

Let $A \in \mathscr{G}_{\mathsf{NW}}$. The goal is to show that $A = D \cdot P$, where $D, P \in \mathscr{G}_{\mathsf{NW}}$ are diagonal and permutation matrices respectively. As A is block-permuted (by Claim 4.1), there is a permutation μ on $[d]$ such that the only non-zero blocks of A are the $(i, \mu(i))$-th blocks for $i \in [d]$. Lemma 2.1 implies,

$$H_{\mathsf{NW}}(\mathbf{x}) = A^T \cdot H_{\mathsf{NW}}(A \cdot \mathbf{x}) \cdot A. \tag{10}$$

The rows and columns of $H_{\mathsf{NW}}(\mathbf{x})$ and $H_{\mathsf{NW}}(A \cdot \mathbf{x})$ are indexed by the $\mathbf{x}$-variables, and the i-th block of rows and columns by the $\mathbf{x}_i$-variables for $i \in [d]$. We can also view $H_{\mathsf{NW}}(\mathbf{x})$ and $H_{\mathsf{NW}}(A \cdot \mathbf{x})$ as block matrices with the (i,l)-th block defined by the i-th block of rows and l-th block of columns. Let C_{il} and B_{il} be the (i,l)-th blocks of $H_{\mathsf{NW}}(\mathbf{x})$ and $H_{\mathsf{NW}}(A \cdot \mathbf{x})$ respectively. Then

$$C_{il} = \left(\frac{\partial^2 \mathsf{NW}}{\partial x_{i,j} \partial x_{l,r}}\right)_{j,r\in[d]} \quad \text{and} \quad B_{il} = \left(\frac{\partial^2 \mathsf{NW}}{\partial x_{i,j} \partial x_{l,r}}(A \cdot \mathbf{x})\right)_{j,r\in[d]}. \tag{11}$$

Observation D.4. *Let $\pi = \mu^{-1}$. Then, for every $i, l \in [d]$,*

$$(A^T_{\pi(i)i})^{-1} \cdot C_{il} \cdot (A_{\pi(l)l})^{-1} = B_{\pi(i)\pi(l)}. \tag{12}$$

Proof. The only non-zero block among the i-th block of rows in A^T is $A^T_{\pi(i)i}$, and the only non-zero block among the l-th block of columns in A is $A_{\pi(l)l}$. Hence, from Equation (10), we have $C_{il} = A^T_{\pi(i)i} \cdot B_{\pi(i)\pi(l)} \cdot A_{\pi(l)l}$. As A is block-permuted and invertible, $A^T_{\pi(i)i}$, $A_{\pi(l)l}$ are also invertible. □

For contradiction, suppose A is not a product of a diagonal matrix and a permutation matrix. As A is block-permuted, there is a $l \in [d]$ such that $A_{\pi(l)l}$ has a column containing more than one non-zero entries which implies $(A_{\pi(l)l})^{-1}$ also has a column containing more than one non-zero entries; let this be the r-th column of $(A_{\pi(l)l})^{-1}$, where $r \in [d]$. We work with this choice of l and r, and fix $i \in [d]\backslash\{l\}$ arbitrarily, in Equation (12). For $j \in [d]$, let g_{jr} and f_{jr} be the (j,r)-th entries of the matrices in the LHS and RHS of Equation (12) respectively. As $g_{jr} = f_{jr}$, the evaluation dimensions of g_{jr} and f_{jr} must be equal with respect to every $\mathbf{z} \subseteq \mathbf{x}$. However, the following claim shows that this is false. Thus, A is a product of a diagonal matrix and a permutation matrix.

Claim D.5. *Let $d \geq 2k+4$. For every $j \in [d]$, there exists $\mathbf{z} \subseteq \mathbf{x}$ such that* $\text{evalDim}_{\mathbf{z}}(g_{jr}) > \text{evalDim}_{\mathbf{z}}(f_{jr})$.

The proof of Claim D.5 is given in Section D.3.1. It is a simple exercise to show that if $A \in \mathscr{G}_{\mathsf{NW}}$ and $A = D \cdot P$, where D and P are diagonal and permutation matrices respectively, then $D, P \in \mathscr{G}_{\mathsf{NW}}$.

D.3.1 Proof of Claim D.5

Recall the choice of l, r and i from the paragraph before the statement of Claim D.5.

Observation D.5. *For every $j, s \in [d]$, the (j,s)-th entry of C_{il} equals*

$$\sum_{\substack{h \in \mathbb{F}_d[z]_k \\ h(i)=j,\, h(l)=s}} \prod_{t \in [d]\backslash\{i,l\}} x_{t,h(t)}.$$

The number of monomials in the above polynomial is d^{k-1}.

Proof. The proof follows directly from Equation (11). □

Observation D.6. *The polynomials in two distinct entries of C_{il} are monomial disjoint.*

Proof. Let $(j,s) \neq (j',s')$. The monomials of the polynomial at the (j,s)-th entry of C_{il} correspond to univariate polynomials $h \in \mathbb{F}_d[z]_k$ such that $h(i) = j$ and $h(l) = s$, whereas the monomials of the polynomial at the (j',s')-th entry of C_{il} correspond to univariate polynomials $h' \in \mathbb{F}_d[z]_k$ such that $h'(i) = j'$ and $h'(l) = s'$. As two distinct degree-k univariates share at most k roots over $\mathbb{F}_d$ and $d - 2 \geq k+1$, the two polynomials must be monomial disjoint. □

Recall that g_{jr} is the (j,r)-th entry of $(A^T_{\pi(i)i})^{-1} \cdot C_{il} \cdot (A_{\pi(l)l})^{-1}$ and f_{jr} is the (j,r)-th entry of $B_{\pi(i)\pi(l)}$.

Observation D.7. *For every $j \in [d]$, $g_{jr} \neq 0$ is a $\mathbb{F}$-linear combination of at least two entries of C_{il}.*

Proof. The proof follows immediately from the choice of l and r, and by observing that none of the rows of $(A^T_{\pi(i)i})^{-1}$ has all zero entries. □

Now, pick an arbitrary set $T \subseteq [d] \setminus \{i,l\}$ such that $|T| = k+1$; this is possible as $d-2 \geq k+1$. Fix $\mathbf{z} = \biguplus_{w \in T} \mathbf{x}_w$.

Observation D.8. *For every $j \in [d]$, $\text{evalDim}_{\mathbf{z}}(f_{jr}) \leq d^{k-1}$.*

Proof. From Equation (11), we have

$$f_{jr} = \frac{\partial^2 \mathsf{NW}}{\partial x_{\pi(i),j} \partial x_{\pi(l),r}}(A \cdot \mathbf{x}).$$

Thus, f_{jr} is computed by a depth three circuit having top fan-in d^{k-1}. Further, as A is block-permuted, the circuit is set-multilinear with respect to the partition $\biguplus_{t \in [d]\setminus\{i,l\}} \mathbf{x}_t$. In other words, f_{jr} can be expressed as a sum of d^{k-1} many products of linear forms such that each product term is of the form $\prod_{t \in [d]\setminus\{i,l\}} \ell_t(\mathbf{x}_t)$, where ℓ_t is a linear form. The proof is immediate from this point. □

Observation D.9. *For every $j \in [d]$, $\text{evalDim}_{\mathbf{z}}(g_{jr}) \geq 2 \cdot d^{k-1}$.*

Proof. From Observations D.5, D.6 and D.7, we can infer that there exists a set $P \subseteq \mathbb{F}_d[z]_k$ of size $|P| \geq 2 \cdot d^{k-1}$ such that

$$g_{jr} = \sum_{h \in P} \beta_h \cdot \prod_{t \in [d]\setminus\{i,l\}} x_{t,h(t)}, \quad \text{where } \beta_h \in \mathbb{F}\setminus\{0\}.$$

Now, we argue that $\text{evalDim}_{\mathbf{z}}(g_{jr}) = |P|$. Clearly, $\text{evalDim}_{\mathbf{z}}(g_{jr}) \leq |P|$. For a fixed $h \in P$ and every $w \in T$, set the variables $x_{w,h(w)} = 1$ and the remaining variables of $\mathbf{z}$ to 0. This substitution reduces the above sum to a single term $\beta_h \cdot \prod_{t \in [d]\setminus(\{i,l\}\uplus T)} x_{t,h(t)}$, as $d-2 \geq k+1$. Moreover,

$$\prod_{t \in [d]\setminus(\{i,l\}\uplus T)} x_{t,h(t)} \neq \prod_{t \in [d]\setminus(\{i,l\}\uplus T)} x_{t,h'(t)},$$

for distinct $h, h' \in P$, as $(d-2)-(k+1) \geq k+1$ (by assumption). Hence, under various similar substitutions of the $\mathbf{z}$-variables, we get $|P|$ distinct monomials implying $\text{evalDim}_{\mathbf{z}}(g_{jr}) \geq |P|$. □

D.4 Proofs from Section 4.3

D.4.1 Proof of Lemma 4.2

Lemma 4.2 (restated): *Let $\mathbb{F}$ be either $\mathbb{R}, \mathbb{Q}$ or finite field such that $d \nmid |\mathbb{F}|-1$. If $D \in \mathscr{G}_{\mathsf{NW}}$ is a diagonal matrix over $\mathbb{F}$ then D is of the form $D = \text{diag}(\beta_0, \ldots, \beta_{d-1}) \otimes I_d$, where each $\beta_i \in \mathbb{F}$ and $\prod_{i \in [d]} \beta_i = 1$.*

Proof. Let $\mathbb{F} = \mathbb{R}$. Let $D \in \mathscr{G}_{\mathsf{NW}}$ be a diagonal matrix with real entries, and the $((i,j),(i,j))$-th entry of D be $\beta_{i,j} \in \mathbb{R}$ for $i,j \in [d]$. We can express $\beta_{i,j}$ as $\beta_{i,j} = (-1)^{\lambda_{i,j}} \cdot 2^{\gamma_{i,j}}$, where $\lambda_{i,j} \in \{0,1\}$ and $\gamma_{i,j} \in \mathbb{R}$. When D is applied on $\mathbf{x}$, a monomial $m_h = \prod_{i \in \mathbb{F}_d} x_{i,h(i)}$ of NW gets mapped to $\left(\prod_{i \in \mathbb{F}_d} (-1)^{\lambda_{i,h(i)}} \cdot 2^{\gamma_{i,h(i)}}\right) \cdot m_h$, implying $\prod_{i \in \mathbb{F}_d} (-1)^{\lambda_{i,h(i)}} = \prod_{i \in \mathbb{F}_d} 2^{\gamma_{i,h(i)}} = 1$. In other words,

$$\begin{aligned} \sum_{i \in [d]} \lambda_{i,h(i)} &= 0 \quad \text{over } \mathbb{F}_2, \text{ for all } h \in \mathbb{F}_d[z]_k, \text{ and} \\ \sum_{i \in [d]} \gamma_{i,h(i)} &= 0 \quad \text{over } \mathbb{R}, \text{ for all } h \in \mathbb{F}_d[z]_k. \end{aligned}$$

By invoking Corollary D.1 (over $\mathbb{F} = \mathbb{F}_2$ and over $\mathbb{F} = \mathbb{R}$) for the above two linear systems, we get $\lambda_{i,0} = \ldots = \lambda_{i,d-1} = \lambda_i$ and $\gamma_{i,0} = \ldots = \gamma_{i,d-1} = \gamma_i$ for every $i \in [d]$, where $\lambda_0, \ldots, \lambda_{d-1} \in \mathbb{F}_2$ (similarly, $\gamma_0, \ldots, \gamma_{d-1} \in \mathbb{R}$) satisfy $\sum_{i\in[d]} \lambda_i = 0$ in $\mathbb{F}_2$ (similarly, $\sum_{i\in[d]} \gamma_i = 0$ in $\mathbb{R}$). This implies $\beta_{i,0} = \ldots = \beta_{i,d-1} = \beta_i$ for every $i \in [d]$, where $\beta_0, \ldots, \beta_{d-1} \in \mathbb{R}$ satisfy $\prod_{i\in[d]} \beta_i = 1$. As $\mathbb{Q}$ is a sub-field of $\mathbb{R}$, NW can not have a diagonal symmetry other that $\mathrm{diag}(\beta_0, \ldots, \beta_{d-1}) \otimes I_d$ over $\mathbb{Q}$.

Let $\mathbb{F}$ be a finite field and $\mathcal{R} = \mathbb{Z}_{|\mathbb{F}|-1}$, such that $d \nmid |\mathbb{F}| - 1$. Let $D = \mathrm{diag}(\beta_{0,0}, \ldots, \beta_{d-1,d-1})$, where $\beta_{i,j} \in \mathbb{F}$ for every $i, j \in [d]$. Then for every $i, j \in [d]$, $\beta_{i,j}$ can be written as $\beta_{i,j} = \tau^{\delta_{i,j}}$, where τ is a generator of $\mathbb{F}^\times$. When D is applied to $\mathbf{x}$, a monomial $m_h = \prod_{i\in\mathbb{F}_d} x_{i,h(i)}$ of NW gets mapped to $(\prod_{i\in\mathbb{F}_d} \tau^{\delta_{i,j}}) \cdot m_h$. As $D \in \mathscr{G}_{\mathsf{NW}}$, $\prod_{i\in\mathbb{F}_d} \tau^{\delta_{i,j}} = 1$, which implies

$$\sum_{i\in[d]} \delta_{i,h(i)} = 0 \quad \text{over } \mathcal{R}, \quad \text{for all } h \in \mathbb{F}_d[z]_k.$$

By invoking Claim D.4 over $\mathcal{R}$ for the above system, we get the desired result. □

Corollary D.3. *The diagonal symmetries of* NW *over* $\mathbb{F}$ *are contained in the group of symmetries of every set-multilinear polynomial over* $\mathbb{F}$.

D.4.2 Proof of Theorem 6

Theorem 6 (restated): *Let* $\mathbb{F}$ *be either* $\mathbb{R}, \mathbb{Q}$ *or a finite field such that* $d \nmid |\mathbb{F}| - 1$. *Then,* NW *is not characterized by its symmetries over* $\mathbb{F}$.

Proof. We know that the symmetries of NW are generated by block-permuted permutation matrices and diagonal matrices (Theorem 5). Let $P_1, \ldots, P_r$ be all the permutation symmetries of NW. We now show that there exists a set-multilinear polynomial $f \in \mathbb{F}[\mathbf{x}]$ such that $f \neq \alpha \cdot \mathsf{NW}$ for any $\alpha \in \mathbb{F}$ but $\mathscr{G}_{\mathsf{NW}} \subseteq \mathscr{G}_f$. Let $h \in \mathbb{F}_d[z]$ of degree $k+1$ and $m_h := \prod_{i\in[d]} x_{i,h(i)}$. Let S be the smallest set of monomials containing m_h such that for every monomial $m \in S$, $m(P_i \cdot \mathbf{x}) \in S$ for every $i \in \{1, \ldots, r\}$. Clearly, S is a set of set-multilinear monomials. Suppose $f \in \mathbb{F}[\mathbf{x}]$ is defined as follows

$$f = \sum_{m\in S} m.$$

As f is a set-multilinear polynomial, by Corollary D.3 all the diagonal symmetries of NW are contained in $\mathscr{G}_f$. By definition, all the permutation symmetries of NW are also contained in $\mathscr{G}_f$. Thus, $\mathscr{G}_{\mathsf{NW}} \subseteq \mathscr{G}_f$ but f is not a scalar multiple of NW. □

E Proof of Theorem 3

Theorem 3 (restated): *Suppose* NW *is not computable by circuits of size* s *over a finite field* $\mathbb{F}$, *where* $|\mathbb{F}| \geq 4 \cdot \delta(s)$. *Then, there exist points* $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{F}^n$, *where* $m = \mathrm{poly}(s)$, *such that for every circuit* $\mathtt{C}$ *over* $\mathbb{F}$ *of size at most* s, *there is an* $\ell \in [m]$ *satisfying* $\mathtt{C}(\mathbf{a}_\ell) \neq \mathsf{NW}(\mathbf{a}_\ell)$. *A set of randomly generated points* $\mathbf{a}_1, \ldots, \mathbf{a}_m \in_r \mathbb{F}^n$ *has this property with high probability. Moreover, black-box derandomization of polynomial identity testing for size-*$(10s)$ *circuits over* $\mathbb{F}$ *using* $\mathrm{poly}(s)$ *field operations implies that the above-mentioned points can be computed deterministically using* $\mathrm{poly}(s)$ *field operations.*

Proof. Let $\mathtt{C}$ be a circuit of size s over a finite field $\mathbb{F}$. As NW is not computable by size-s circuits over $\mathbb{F}$ (by assumption), $\mathtt{C}(\mathbf{x}) - \mathsf{NW} \neq 0$. The polynomial $\mathtt{C}(\mathbf{x}) - \mathsf{NW}$ has degree bounded by $\delta(s)$, as $\delta(s) \geq d$. By Schwartz-Zippel lemma, for any $m \in \mathbb{N}$,

$$\Pr_{\mathbf{a}_1,\ldots,\mathbf{a}_m \in_r \mathbb{F}^n} [\mathtt{C}(\mathbf{a}_\ell) = \mathsf{NW}(\mathbf{a}_\ell), \text{ for all } \ell \in [m]] \leq \left(\frac{\delta(s)}{|\mathbb{F}|}\right)^m.$$

The number of size-s circuits over $\mathbb{F}$ is at most $2^{s^2+s} \cdot |\mathbb{F}|^s$ (as there are 2^s ways to label the nodes as $+$ and $\times$ gates, at most 2^{s^2} ways to choose the adjacency matrix of the underlying directed graph, and $|\mathbb{F}|^s$ ways to label the edges of a given graph). Therefore,

$$\Pr_{\mathbf{a}_1,\ldots,\mathbf{a}_m \in_r \mathbb{F}^n} [\exists \text{ a size-}s \text{ circuit } \mathtt{C} \text{ such that } \mathtt{C}(\mathbf{a}_\ell) = \mathsf{NW}(\mathbf{a}_\ell), \text{ for all } \ell \in [m]] \leq |\mathbb{F}|^s \cdot 2^{s^2+s} \cdot \left(\frac{\delta(s)}{|\mathbb{F}|}\right)^m.$$

By fixing $m = s^2 + 2s$, the above probability can be upper bounded by $\exp(-s)$ as $|\mathbb{F}| \geq 4 \cdot \delta(s)$.

Now, let us show that black-box derandomization of identity testing implies that such points $\mathbf{a}_1, \ldots, \mathbf{a}_m$ can be computed deterministically. Consider the class $\mathcal{C}$ of size-$(10s)$ circuits over $\mathbb{F}$ on $n+1$ variables $\mathbf{x} \uplus u$. Assume that $\mathcal{H} = \{(\mathbf{b}_0, \mu_0), \ldots, (\mathbf{b}_{w-1}, \mu_{w-1})\} \subseteq \mathbb{F}^{n+1}$ is a hitting set[15] for the circuit class $\mathcal{C}$, and $\mathcal{H}$ is computable using $\mathrm{poly}(s)$ field operations. Let $\mathcal{P} \subseteq \mathbb{F}^n$ be the set of points that includes $\mathbf{b}_0, \ldots \mathbf{b}_{w-1}$ along with $A_i(\mu_\ell) \cdot \mathbf{b}_\ell$, $A_{a,r} \cdot \mathbf{b}_\ell$ and $A_t \cdot \mathbf{b}_\ell$ for every $\ell \in [w], i \in [d], a \in \mathbb{F}_d^\times, r \in [k+1]$ and $t \in [d] \backslash [k+1]$. Finally, $\mathcal{P}$ also contains the point $\mathbf{b} \in \mathbb{F}^n$ obtained by setting $x_{i0} = 1$, for $i \in [d]$, and all other variables to zero. Observe that $|\mathcal{P}| = \mathrm{poly}(s)$ as $|\mathcal{H}| = \mathrm{poly}(s)$.

Claim E.1. *For every size-s circuit $\mathtt{C}$ on n inputs, there is a point $\mathbf{a}$ in $\mathcal{P}$ such that $\mathtt{C}(\mathbf{a}) \neq \mathsf{NW}(\mathbf{a})$.*

Proof. As NW is not computable by size-s circuits, $f = \mathtt{C}(\mathbf{x}) \neq \alpha \cdot \mathsf{NW}$ for all $\alpha \in \mathbb{F}^\times$ [16]. Hence, at least one of the identities, in the proof of Lemma 3.1, is not satisfied by f unless $f = 0$. If $f = 0$ then $f(\mathbf{b}) \neq \mathsf{NW}(\mathbf{b}) = 1$, and so let $f \neq 0$. The degrees of the polynomials $q_1(f(A_i(u) \cdot \mathbf{x}), f(\mathbf{x}), u)$, $q_2(f(A_{a,r} \cdot \mathbf{x}), f(\mathbf{x}))$ and $q_3(f(A_t \cdot \mathbf{x}))$ are upper bounded by $2 \cdot \delta(s)$. Also, it can be verified that the polynomials $q_1(f(A_i(u) \cdot \mathbf{x}), f(\mathbf{x}), u)$, $q_2(f(A_{a,r} \cdot \mathbf{x}), f(\mathbf{x}))$ and $q_3(f(A_t \cdot \mathbf{x}))$ are computable by size-$(10s)$ circuits on $n+1$ variables $\mathbf{x} \uplus u$. Hence, $\mathcal{H}$ is a hitting-set for these polynomials. Without loss of generality, let $q_1(f(A_i(u) \cdot \mathbf{x}), f(\mathbf{x}), u) = 0$ be an identity that is not satisfied by f. Then, there is a $(\mathbf{b}_\ell, \mu_\ell) \in \mathcal{H}$ such that $q_1(f(A_i(\mu_\ell) \cdot \mathbf{b}_\ell), f(\mathbf{b}_\ell), \mu_\ell) \neq 0$ implying $f(A_i(\mu_\ell) \cdot \mathbf{b}_\ell) \neq \mu_\ell \cdot f(\mathbf{b}_\ell)$. On the other hand, $\mathsf{NW}(A_i(\mu_\ell) \cdot \mathbf{b}_\ell) = \mu_\ell \cdot \mathsf{NW}(\mathbf{b}_\ell)$ as NW satisfies all the identities. Therefore, either $f(A_i(\mu_\ell) \cdot \mathbf{b}_\ell) \neq \mathsf{NW}(A_i(\mu_\ell) \cdot \mathbf{b}_\ell)$ or $f(\mathbf{b}_\ell) \neq \mathsf{NW}(\mathbf{b}_\ell)$. This implies the claim as $A_i(\mu_\ell) \cdot \mathbf{b}_\ell$ and $\mathbf{b}_\ell$ belong to $\mathcal{P}$. □

The proof of the theorem follows from the above claim and by observing that $\mathcal{P}$ can be constructed from $\mathcal{H}$ using $\mathrm{poly}(s)$ field operations.

□

[15] A set of points $\mathcal{H}$ is a hitting-set for a circuit class $\mathcal{C}$ if for every circuit $\mathtt{C} \in \mathcal{C}$ computing a non-zero polynomial, there exists a point $\mathbf{b} \in \mathcal{H}$ such that $\mathtt{C}(\mathbf{b}) \neq 0$. Black-box derandomization of identity testing for a circuit class amounts to constructing a hitting-set for the class.

[16] If $\alpha \cdot \mathsf{NW}$ is computable by a size-s circuit $\mathtt{C}$, for some $\alpha \in \mathbb{F}^\times$, then NW is also computable by a size-s circuit by appropriately scaling some of the edges feeding into the output gate of $\mathtt{C}$ by α^{-1}.

F Proofs from Section 6

F.1 Proof of Claim 6.1

Claim 6.1 (restated): *With high probability, matrix D can be computed in* $\text{poly}(d, \rho)$ *time. Moreover, f is equivalent to* NW *if and only if $f(D \cdot \mathbf{x})$ is BP equivalent to* NW.

Proof. Suppose f is equivalent to NW, i.e., $f = \text{NW}(A \cdot \mathbf{x})$ for $A \in \text{GL}_{d^2}(\mathbb{F})$. Then, $R_1, \ldots, R_{d-1}$ is a basis of $\mathfrak{g}_{\text{NW}}$, where $L_i = A^{-1} \cdot R_i \cdot A$ (Lemma 2.2). We know that $L = a_1 L_1 + \cdots + a_{d-1} L_{d-1}$, where $a_1, \ldots, a_{d-1}$ are chosen uniformly at random from S. Pretend that $\mathbf{a} = \{a_1, \ldots, a_{d-1}\}$ are formal variables. Then, $L = A^{-1} \cdot R \cdot A$, where $R = a_1 R_1 + \cdots + a_{d-1} R_{d-1}$. It is easy to see that $R = \text{diag}(\alpha_1, \ldots, \alpha_d) \otimes I_d$, where $\alpha_1, \ldots, \alpha_d$ are linear forms in $\mathbf{a}$-variables, and $\alpha_d = -(\sum_{i=1}^{d-1} \alpha_i)$. By Lemma 4.1 and as $|\mathbb{F}| \geq d^2$, there is a setting of the $\mathbf{a}$-variables that makes $\alpha_1, \ldots, \alpha_d$ distinct field elements. In other words, $\alpha_1, \ldots, \alpha_d$ are pairwise distinct linear forms in $\mathbf{a}$-variables. Hence, from the Schwartz-Zippel lemma [Zip79,Sch80], on setting $a_1, \ldots, a_d$ uniformly at random from S, $\alpha_1, \ldots, \alpha_d$ become distinct elements of $\mathbb{F}$ with high probability.

Compute the characteristic polynomial of L, denoted $h_L(z)$ and factorize it. As f is equivalent to NW, L and R are similar matrices and their characteristic polynomials are the same. Then $h_L(z)$ factorizes as $h_L(z) = (z - \beta_1)^d \cdots (z - \beta_d)^d$, for distinct $\beta_1, \ldots, \beta_d \in \mathbb{F}$ such that there is an (unknown) permutation σ on $[d]$ such that $\beta_i = \alpha_{\sigma(i)}$ for $i \in [d]$. Suppose $B = \text{diag}(\beta_1, \ldots, \beta_d) \otimes I_d$. Let D be a $d^2 \times d^2$ size formal matrix such that

$$L \cdot D = D \cdot B. \tag{13}$$

Solve the system of linear equations obtained from Equation (13) (by treating the entries of D as variables) and pick a random matrix from the solution space; call this solution matrix D. With high probability D is invertible (as $D = A^{-1}P$ is also in the solution space for a suitable permutation matrix P). Equation (13) implies that

$$R \cdot A \cdot D = A \cdot D \cdot B.$$

Recall that $R = \text{diag}(\alpha_1, \ldots, \alpha_d) \otimes I_d$ and $B = \text{diag}(\alpha_{\sigma(1)}, \ldots, \alpha_{\sigma(d)}) \otimes I_d$. As $\alpha_1, \ldots, \alpha_d$ are distinct, it is an easy exercise to show that AD is a block permuted matrix. Hence $f(D \cdot \mathbf{x})$ is BP equivalent to NW. □

F.2 Proofs from Section 6.1.1

F.2.1 Proof of Claim 6.2

Claim 6.2 (restated): *Suppose $f \in \mathbb{F}[\mathbf{x}]$ is BD permutation equivalent to* NW. *Then, there exist permutations $\sigma_0, \ldots, \sigma_{d-1}$ on $[d]$ such that $\sigma_0(0) = \cdots = \sigma_k(0) = 0, \sigma_0(1) = 1$ and $A = \text{diag}(M_{\sigma_0}, \ldots, M_{\sigma_{d-1}})$ satisfies $f = \text{NW}(A \cdot \mathbf{x})$.*

Proof. Since $f \in \mathbb{F}[\mathbf{x}]$ is BD permutation equivalent to NW, there exist permutations $\pi_0, \ldots, \pi_{d-1}$ on $[d]$, such that $A' = \text{diag}(M_{\pi_0}, \ldots, M_{\pi_{d-1}})$ satisfies $f = \text{NW}(A' \cdot \mathbf{x})$. Let $h \in \mathbb{F}_d[z]_k$ such that $\pi_0(0) = h(0), \ldots, \pi_k(0) = h(k)$. For $i \in [d]$, define $\sigma_i : \mathbb{F}_d \to \mathbb{F}_d$ as

$$\sigma_i(l) := \alpha \cdot (\pi_i(l) - h(i)) \text{ for all } l \in [d],$$

where $\alpha := \frac{1}{\pi_0(1)-h(0)}$. Note that for every $i \in [d]$, σ_i is well defined as $\pi_0(1) \neq h(0)$. The following observation can be verified easily.

Observation F.1. *$\sigma_0, \ldots, \sigma_{d-1}$ are permutations on $\mathbb{F}_d$. Also, $\sigma_0(0) = \cdots = \sigma_k(0) = 0$ and $\sigma_0(1) = 1$.*

For $i \in [d]$, let $\tau_i : \mathbb{F}_d \to \mathbb{F}_d$ be defined as $\tau_i(l) := \alpha \cdot (l - h(i))$ for every $l \in \mathbb{F}_d$. Observe that $\tau_0, \ldots, \tau_{d-1}$ are permutations on $\mathbb{F}_d$ and for every $i \in [d]$

$$\sigma_i = \tau_i \circ \pi_i. \tag{14}$$

Let $A = \mathrm{diag}(M_{\sigma_0}, \ldots, M_{\sigma_{d-1}}), C = \mathrm{diag}(M_{\tau_0}, \ldots, M_{\tau_{d-1}})$. As A, A', C are block diagonal matrices, the above equation implies

$$A = C \cdot A'.$$

Observation F.2. $C \in \mathscr{G}_{\mathsf{NW}}$.

Proof. On applying C on $\mathbf{x}$, $x_{i,j}$ gets mapped to $x_{i,\alpha\cdot(j-h(i))}$ for every $i, j \in [d]$. This shows $C \in \mathscr{G}_{\mathsf{NW}}$ (similar to item 3 of Claim 3.1). □

Since $\mathsf{NW}(\mathbf{x}) = \mathsf{NW}(C \cdot \mathbf{x})$, we get $f = \mathsf{NW}(C \cdot A' \cdot \mathbf{x}) = \mathsf{NW}(A \cdot \mathbf{x})$. □

F.2.2 Proof of Claim 6.3

Claim 6.3 (restated): *The list of $d - k$ nice polynomials $\{h_0, \ldots, h_{d-k-1}\}$ can be computed in $\mathrm{poly}(d, \rho)$ time.*

Proof. We create two lists of $d - k$ distinct polynomials in $\mathbb{F}_d[z]_k$, namely the p-list and the h-list as described below. Then we show that the h-list is a list of nice polynomials.

Procedure to create h-list and p-list:

1. Interpolate $(0,0), \ldots, (k,0)$ to get $p_0 \in \mathbb{F}_d[z]_k$ and then interpolate $(0,1), (1,0), \ldots, (k-1,0), (k,0)$ to get $h_0 \in \mathbb{F}_d[z]_k$. (In this case, $p_0 = 0$ and $h_0 \neq 0$.)
2. Interpolate $(0,0), \ldots, (k-1,0), (k+1, h_0(k+1))$ to get $p_1 \in \mathbb{F}_d[z]_k$ and then interpolate $(0,1), (1,0) \ldots, (k-1,0), (k, p_1(k))$ to get $h_1 \in \mathbb{F}_d[z]_k$.
3. For $r \in \{2, \ldots, d-k-1\}$ do the following.
   (a) For $r_1 = 1$ to r, interpolate $(0,0), \ldots, (k-1,0), (k+r_1, h_{r-1}(k+r_1))$ to get $\tilde{p}_{r_1} \in \mathbb{F}_d[z]_k$. (It is argued in Observation F.4 that $\tilde{p}_1, \ldots, \tilde{p}_r$ are distinct polynomials.) Pick a polynomial from $\tilde{p}_1, \ldots, \tilde{p}_r$ that is different from each of $p_0, \ldots, p_{r-2}$. Set that polynomial to be p_r. (It is argued in Observation F.5 that no polynomial amongst $\tilde{p}_1, \ldots, \tilde{p}_r$ is equal to p_{r-1}, and so $p_r \neq p_i$ for all $i \in [r]$.)
   (b) Interpolate $(0,1), (1,0), \ldots, (k-1,0), (k, p_r(k))$ to get $h_r \in \mathbb{F}_d[z]_k$.

We note some easy-to-verify observations about these lists.

Observation F.3. *1. The p-list and h-list can be computed in $\mathrm{poly}(d)$ time and they do not have a polynomial in common.*

2. *All polynomials in the p-list (similarly in the h-list) agree on k points, namely* $0, \ldots, k-1$.

3. *For distinct* $r, r' \in [d-k]$, p_r *and* h_r *agree on k points* $1, \ldots, k$, *and* p_r *and* $h_{r'}$ *agree on* $k-1$ *points* $1, \ldots, k-1$.

The following two sub claims imply that $\{h_0, \ldots, h_{d-k-1}\}$ is a list of $d-k$ distinct nice polynomials.

Subclaim F.1. *Each of the p-list and h-list contains* $d-k$ *distinct polynomials.*

Its proof is given in Section F.2.3. The following fact would be required to prove that $\{h_0, \ldots, h_{d-k-1}\}$ is a list of nice polynomials.

Fact 2. *Suppose* $h \in \mathbb{F}_d[z]_k$ *and* $i_0, \ldots, i_k \in [d]$ *be distinct elements. Then, given* $\sigma_{i_0}(h(i_0)), \ldots, \sigma_{i_k}(h(i_k))$, *we can compute* $\sigma_i(h(i))$ *for every* $i \in [d] \setminus \{i_0, \ldots, i_k\}$ *in* $\mathrm{poly}(d, \rho)$ *time.*

The proofs of Fact 2 and Subclaim F.2 are also given in Section F.2.3.

Subclaim F.2. *For every* $r \in [d-k]$, $\sigma_0(h_r(0)), \ldots, \sigma_k(h_r(k))$ *can be computed in* $\mathrm{poly}(d, \rho)$ *time.*

□

F.2.3 Proofs of subclaims and fact used in Claim 6.3

Proof of Subclaim F.1

Proof. For some $r \in [d-k]$, item 2 of Observation F.3 implies that if $p_0, \ldots, p_r$ are pairwise distinct then $h_0, \ldots, h_r$ are also pairwise distinct. We show that $p_0, \ldots, p_r$ are pairwise distinct polynomials by induction on r. The base case, i.e. $r = 0$ is trivially satisfied. Suppose the hypothesis holds for $r-1$, i.e. $p_0, \ldots, p_{r-1}$ are pairwise distinct. This implies $h_0, \ldots, h_{r-1}$ are also pairwise distinct. We construct r polynomials $\tilde{p}_1, \ldots \tilde{p}_r$ in $\mathbb{F}_d[z]_k$ by interpolating $(0,0), \ldots, (k-1,0), (k+1, h_{r-1}(k+1))$; $\ldots$; $(0,0), \ldots, (k-1,0), (k+r, h_{r-1}(k+r))$ respectively.

Observation F.4. $\tilde{p}_1, \ldots, \tilde{p}_r$ *are distinct polynomials in* $\mathbb{F}_d[z]_k$.

Proof. Suppose not, then there exist distinct $r_1, r_2 \in \{1, \ldots, r\}$, such that $\tilde{p}_{r_1} = \tilde{p}_{r_2}$. This implies that the polynomials $\tilde{p}_{r_1}$ and h_{r-1} agree on $k+1$ points $1, \ldots, k-1, k+r_1$ and $k+r_2$, which is a contradiction as $\tilde{p}_{r_1}$ and h_{r-1} are distinct polynomials (as $\tilde{p}_{r_1}(0) = 0$ whereas $h_{r-1}(0) = 1$). □

Observation F.5. *For every* $r_1 \in \{1, \ldots, r\}$, $\tilde{p}_{r_1} \neq p_{r-1}$.

Proof. Suppose not. Then, there exists $r_1 \in \{1, \ldots, r\}$ such that, $\tilde{p}_{r_1} = p_{r-1}$. Then, $p_{r-1}(k+r_1) = \tilde{p}_{r_1}(k+r_1) = h_{r-1}(k+r_1)$, which along with item 3 of Observation F.3 implies that h_{r-1} and p_{r-1} agree on $k+1$ points $1, \ldots, k, k+r_1$, which can not happen as p_{r-1} and h_{r-1} are distinct polynomials. □

Hence $p_0, \ldots, p_r$ are distinct polynomials. □

Proof of Fact 2:

Proof. Since f is block-diagonal permutation equivalent to NW, Observation 6.1 implies that on setting $x_{i_0,\sigma_{i_0}(h(i_0))} = \cdots = x_{i_k,\sigma_{i_k}(h(i_k))} = 1$ and other variables of $\mathbf{x}_{i_0}, \ldots, \mathbf{x}_{i_k}$ equal to zero, f reduces to

$$c \cdot \prod_{i \in [d] \setminus \{i_0,\ldots,i_k\}} x_{i,\sigma_i(h(i))}, \qquad \text{where } c \in \mathbb{F}.$$

It is easy to show that in this case $\sigma_i(h(i))$ for $i \in [d] \setminus \{i_0, \ldots, i_k\}$ can be recovered in $\text{poly}(d, \rho)$ time from black-box access to f. □

Proof of Subclaim F.2

Proof. For every $r \in [d-k], \sigma_0(h_r(0)) = 1, \sigma_1(h_r(1)) = \cdots = \sigma_{k-1}(h_r(k-1)) = 0$ from Step 1 of Algorithm 3. We show that $\sigma_k(h_r(k))$ can be computed efficiently by induction on r. When $r = 0$, we know that $\sigma_k(h_0(k)) = \sigma_k(0) = 0$. Thus, the base case holds. Suppose that the hypothesis holds for $r-1$, i.e. we can efficiently compute $\sigma_k(h_{r-1}(k))$. Recall that p_r is computed by interpolating $(0,0), \ldots, (k-1,0), (k+r_1, h_{r-1}(k+r_1))$ for some $r_1 \in \{1, \ldots, r\}$. Using Fact 2 on $\sigma_0(h_{r-1}(0)), \ldots, \sigma_{k-1}(h_{r-1}(k-1)), \sigma_k(h_{r-1}(k))$ we compute $\sigma_{k+r_1}(h_{r-1}(k+r_1)) = \sigma_{k+r_1}(p_r(k+r_1))$ and then using Fact 2 again on $\sigma_0(0), \ldots, \sigma_{k-1}(0), \sigma_{k+r_1}(p_r(k+r_1))$, we compute $\sigma_k(p_r(k))$, which is equal to $\sigma_k(h_r(k))$. □

F.2.4 Proof of Claim 6.4

Claim 6.4 (restated): *Given a list on nice polynomials $\{h_0, \ldots, h_{d-k-1}\}$, we can recover $d-k$ distinct entries in each of $\sigma_0, \ldots, \sigma_{d-1}$ in $\text{poly}(d, \rho)$ time.*

Proof. We first show that using $\{h_0, \ldots, h_{d-k-1}\}$, we can recover $(d-k)$ distinct entries of each of the permutations $\sigma_{k+1}, \ldots, \sigma_{d-1}$. Fix an $i \in \{k+1, \ldots, d-1\}$. As $h_0, \ldots, h_{d-k-1}$ are nice polynomials, for every $h \in \{h_0, \ldots, h_{d-k-1}\}$, $\sigma_0(h(0)), \ldots, \sigma_k(h(k))$ can be computed efficiently. By invoking Fact 2 on $\sigma_0(h(0)), \ldots, \sigma_k(h(k))$ for every such h, we get $\sigma_i(h_0(i)), \ldots, \sigma_i(h_{d-k-1}(i))$. From item 2 of Observation F.3 and Subclaim F.1 and the fact that σ_i is a permutation, $\sigma_i(h_0(i)), \ldots, \sigma_i(h_{d-k-1}(i))$ are $d-k$ distinct entries of σ_i.

Now using the $d-k$ known entries of σ_{k+1}, we recover $d-k$ distinct entries of each of $\sigma_0, \ldots, \sigma_k$. Suppose there exist distinct $l_0, \ldots, l_{d-k-1} \in [d]$, such that $\sigma_{k+1}(l_0), \ldots, \sigma_{k+1}(l_{d-k-1})$ are known. Fix an $i \in [k+1]$. For $s \in [d-k]$, let p_s be a polynomial in $\mathbb{F}_d[z]_k$ obtained by interpolating $(i', 0), (k+1, l_s)$ for $i' \in [k+1] \setminus \{i\}$. Observe that these are $d-k$ distinct polynomials. Further, for $s_1 \neq s_2$, p_{s_1} and p_{s_2} agree on k points $i' \in [k+1] \setminus \{i\}$ and $p_{s_1}(i) \neq p_{s_2}(i)$, which implies that $(\sigma_i(p_0(i)), \ldots, \sigma_i(p_{d-k-1}(i)))$ is a tuple of distinct entries. Using Fact 2 on $\sigma_{i'}(p_s(i')), \sigma_{k+1}(p_s(k+1))$ for $i' \in [k+1] \setminus \{i\}$, we obtain $d-k$ distinct values $\sigma_i(p_s(i))$ for every $s \in [d-k]$. This shows that for every $i \in [k+1]$, we can compute $d-k$ distinct entries of σ_i efficiently. □

F.2.5 Proof of Claim 6.5

Claim 6.5 (restated): *Suppose $k \in [1, \frac{d}{3}]$. Then, there exist $k+1$ rows in N such that in each of these rows at least $k+1$ entries are known.*

Proof. Suppose this is not true. Then, N has at most k rows such that in each row at least $k+1$ entries are known, and in the remaining at least $d-k$ rows at most k entries are known. This implies that at most $d \cdot k + (d-k)k$ entries are known in N. We know exactly $d(d-k)$ entries in N. Thus, $d(d-k) \leq 2dk - k^2$, which implies $k > \frac{d}{3}$. This is a contradiction. □

F.2.6 Proof of Claim 6.6

Claim 6.6 (restated): *Using $k+1$ rows of N indexed by $l_0, \ldots, l_k$ (as mentioned in Step 4), we can recover all the entries of N in $\text{poly}(d, \rho)$ time.*

Proof. First we show how to recover all the entries of the rows of N indexed by $l_0, \ldots, l_k$. Given that in the rows of N indexed by $l_0, \ldots, l_k$, at least $k+1$ entries are known. For $l \in \{l_0, \ldots, l_k\}$, there exist distinct $i_0, \ldots, i_k \in [d]$, such that $\sigma_{i_0}(l), \ldots, \sigma_{i_k}(l)$ are known. Using Fact 2 on $\sigma_{i_0}(l), \ldots, \sigma_{i_k}(l)$, we recover $\sigma_i(l)$ for every $i \in [d] \setminus \{i_0, \ldots, i_k\}$.

Now we show how to recover $\sigma_i(l)$ for every $l \in [d] \setminus \{l_0, \ldots, l_k\}$ and $i \in [d]$. Let $h = z + (l - i)$. Clearly, $h(i) = l$. Let $i_0, \ldots, i_k \in [d]$ be such that $l_0 = i_0 + l - i, \ldots, l_k = i_k + l - i$. Then, $h(i_0) = l_0, \ldots, h(i_k) = l_k$. Use Fact 2 on the points $\sigma_{i_0}(h(i_0)), \ldots, \sigma_{i_k}(h(i_k))$ to recover $\sigma_i(h(i))$, which is $\sigma_i(l)$. Thus, we recover all the entries of N.

□

F.3 Proofs from Section 6.1.2

F.3.1 Proof of Claim 6.7

Claim 6.7 (restated): *We can assume that $\alpha_{1,0} = \ldots = \alpha_{d-1,0} = 1$ without loss of generality.*

Proof. As f is scaling equivalent to NW, there exists $C = \text{diag}(\beta_{0,0}, \ldots, \beta_{d-1,d-1})$ such that $f = \mathsf{NW}(C \cdot \mathbf{x})$. Suppose $D = \text{diag}(a, \beta_{1,0}^{-1}, \ldots \beta_{d-1,0}^{-1}) \otimes I_d$, where $a = \prod_{i=1}^{d-1} \beta_{i,0}$. Then, from Claim 3.1, $D \in \mathscr{G}_{\mathsf{NW}}$, which implies $f = \mathsf{NW}(D \cdot C \cdot \mathbf{x})$. Set $B = D \cdot C$. Hence $\alpha_{1,0} = \ldots = \alpha_{d-1,0} = 1$. □

F.3.2 Proof of Claim 6.8

Claim 6.8 (restated): *In Step 4 of Algorithm 4, $\alpha_{i,j}$ can be computed in $\text{poly}(d, \rho)$ time. Further, $f = \mathsf{NW}(B \cdot \mathbf{x})$.*

Proof. For $i, j \in [d]$, suppose $\alpha_{i,j} = \tau^{y_{i,j}}$, where τ is a generator of $\mathbb{F}^\times$. Claim 6.7 implies that $y_{1,0} = \ldots = y_{d-1,0} = 0$. If $f = \mathsf{NW}(B \cdot \mathbf{x})$, then a monomial $m_h = \prod_{i \in \mathbb{F}_d} x_{i,h(i)}$ of NW gets mapped to $c_h \cdot m_h$, where $c_h = \prod_{i \in \mathbb{F}_d} \alpha_{i,h(i)}$. Let $c_h = \tau^{e_h}$. Then, we get the following system of linear equations for every $h \in \mathbb{F}_d[z]_k$ over the ring $\mathbb{Z}_{|\mathbb{F}|-1}$.

$$\sum_{i \in \mathbb{F}_d} y_{i,h(i)} = e_h. \tag{15}$$

Recall C, S and $\mathbf{y}$ from Step 3 of the algorithm. On restricting to the polynomials in S, we get the following

$$C \cdot \mathbf{y}^T = \boldsymbol{e},$$

where $\boldsymbol{e} = (e_0\ e_z \ldots e_{(d-1)z}\ e_1\ e_{z+1} \ldots e_{(d-1)z+1} \ldots e_{d-2}\ e_{z+d-2} \ldots\ e_{(d-1)z+d-2}\ e_{d-1})^T$. Recall γ and δ_h from Step 3 and 4. From Cramer's rule, we get

$$y_{i,j} = \gamma \cdot (\sum_{h \in S} e_h \cdot \delta_h) \mod (|\mathbb{F}| - 1),$$

This immediately implies,

$$\alpha_{i,j} = \tau^{y_{i,j}} = \tau^{\gamma \cdot (\sum_{h \in S} e_h \cdot \delta_h) \mod (|\mathbb{F}|-1)},$$

As $c_h = \tau^{e_h}$,

$$\alpha_{i,j} = \prod_{h \in S} c_h^{(\delta_h \cdot \gamma) \mod (|\mathbb{F}|-1)}.$$

As C is a 0/1 matrix, $|\det(C)|$ is bounded by $(d^2 - d + 1)!$, which implies the bit complexity of $\det(C)$ is $\mathrm{poly}(d)$. This implies that the above calculations can be done in $\mathrm{poly}(d, \rho)$ time using repeated squaring. □

F.4 Scaling equivalence test for NW over $\mathbb{R}$

We first state the model of computation over $\mathbb{R}$. We assume that addition, subtraction, multiplication and division of two real numbers can be done in unit time. In addition, we also assume that the positive real root of a univariate real polynomial $y^r - \delta$ can be computed in $\mathrm{poly}(\log r)$ time (see [Bre76, Ye94]).

Suppose a degree d polynomial $f \in \mathbb{F}[\mathbf{x}]$ is scaling equivalent to NW. We wish to find a $B = \mathrm{diag}(\alpha_{0,0}, \ldots, \alpha_{d-1,d-1}) \in \mathrm{GL}_{d^2}(\mathbb{R})$, such that $f = \mathsf{NW}(B \cdot \mathbf{x})$. Note that every $\alpha_{i,j}$ can be written as $\alpha_{i,j} = (-1)^{s_{i,j}} \cdot 2^{\beta_{i,j}}$, where $s_{i,j} \in \mathbb{F}_2$ and $\beta_{i,j} \in \mathbb{R}$. Assume $s_{i,j}, \beta_{i,j}, i, j \in [d]$ are formal variables. Here also, we can assume without loss of generality that $\alpha_{1,0} = \ldots = \alpha_{d-1,0} = 1$, which sets $s_{i,0} = \beta_{i,0} = 0$ for $i \in \{1, \ldots, d-1\}$. For $h \in \{az + b : a \in [d], b \in [d-1]\} \cup \{d-1\}$, let $c_h = (-1)^{\delta_h} \cdot 2^{\gamma_h}$ be the coefficient of $\prod_{i \in \mathbb{F}_d} x_{i,h(i)}$ in f. This gives us the following system of linear equations in $\boldsymbol{\beta}$ and $\boldsymbol{s}$ variables over $\mathbb{R}$ and $\mathbb{F}_2$ respectively.

$$\sum_{i \in \mathbb{F}_d} \beta_{i,h(i)} = \gamma_h \quad \text{and} \quad \sum_{i \in \mathbb{F}_d} s_{i,h(i)} = \delta_h. \tag{16}$$

Hereon, the algorithm for the scaling equivalence test for NW over $\mathbb{R}$ can be obtained by easily adapting Algorithm 4 to solve the system of linear equations mentioned in Equation (16) and compute B.

G Permutation symmetries of NW

Let $P \in \mathscr{G}_{\mathsf{NW}}$ be a permutation matrix, and μ be the corresponding permutation on $\mathbf{x}$, i.e. $\mu(x_{i,j}) = (P \cdot \mathbf{x})(i,j)$. As P is block-permuted (by Claim 4.1), there exist a permutation σ on $\mathbb{F}_d$ and a permutation ϕ on $\mathbb{F}_d[z]_k$ such that for every $h \in \mathbb{F}_d[z]_k$, $\mu(x_{i,h(i)}) = x_{\sigma(i),\phi(h)(\sigma(i))}$. It can be easily verified that a (σ, ϕ) pair (where σ is a permutation on $\mathbb{F}_d$ and ϕ is a permutation on $\mathbb{F}_d[z]_k$) *yields*

a permutation symmetry of NW via the map $\sigma : x_{i,h(i)} \mapsto x_{\sigma(i),\phi(h)(\sigma(i))}$ if and only if for every $i \in \mathbb{F}_d$ and $h_1, h_2 \in \mathbb{F}_d[z]_k$, the following is satisfied:

$$h_1(i) = h_2(i) \quad \text{implies} \quad \phi(h_1)(\sigma(i)) = \phi(h_2)(\sigma(i)) \quad \text{and vice versa.} \tag{17}$$

The task now boils down to understanding which (σ, ϕ) pairs satisfy the above condition. We give a partial answer to this. In particular, we characterize all (σ, ϕ) pair when σ is linear, i.e. there exist $b_\sigma \in \mathbb{F}_d^\times, c_\sigma \in \mathbb{F}_d$, such that for every $i \in \mathbb{F}_d, \sigma(i) = b_\sigma \cdot i + c_\sigma$.

Lemma G.1. *Let $d \geq 4k+1$, σ be a linear permutation on $[d]$ and ϕ be a permutation on $\mathbb{F}_d[z]_k$, such that (σ, ϕ) gives a symmetry of* NW. *Then, there exist $a \in \mathbb{F}_d^\times$ and $p \in \mathbb{F}_d[z]_k$, such that for every $h \in \mathbb{F}_d[z]_k$,*

$$\phi(h) = a \cdot h' + p, \tag{18}$$

where $h' \in \mathbb{F}_d[z]_k$, such that $h'(\sigma(i)) = h(i)$ for every $i \in \mathbb{F}_d$.

The proof of Lemma G.1 is given in Section G.1. It is open if these are the only permutation symmetries of NW.

G.1 Proof of Lemma G.1

Proof. It is easy to check that the pair (σ, ϕ) given in the lemma satisfies Equation (17) and thus gives a symmetry of NW. Now we show that if ϕ is a permutation on $\mathbb{F}_d[z]_k$, such that (σ, ϕ) gives a symmetry of NW, then ϕ satisfies Equation (18) for every $h \in \mathbb{F}_d[z]_k$. We can assume without loss of generality that $\phi(0) = 0$. This is so because if ϕ' is a permutation on $\mathbb{F}_d[z]_k$ defined as $\phi'(h) = \phi(h) - \phi(0)$ for every $h \in \mathbb{F}_d[z]_k$, then (σ, ϕ') gives a symmetry of NW if and only if (σ, ϕ) is a symmetry giving pair. Suppose $\psi : \mathbb{F}_d^\times \to \mathbb{F}_d[z]_k$, defined as $\psi(b) = \phi(b) - \phi(b-1)$ for every non zero constant polynomial $b \in \mathbb{F}_d[z]_k$. Clearly, $\psi(1) = \phi(1)$. We here note some easy to verify observations about ψ.

Observation G.1. *Let $b \in \mathbb{F}_d^\times$. Then, for every $\ell \in \mathbb{F}_d$, $\psi(b)(\ell) \neq 0$.*

Observation G.2. *Suppose $b \in \mathbb{F}_d^\times$. Then, $\phi(b) = \sum_{l=1}^{b} \psi(l)$.*

The following observation is implied from Equation (17) as $h(i) = (h(i))(i)$ for every $i \in \mathbb{F}_d$.

Observation G.3. *Let $h \in \mathbb{F}_d[z]_k$. Then, for every $i \in \mathbb{F}_d$*

$$\phi(h)(\sigma(i)) = \phi(h(i))(\sigma(i)).$$

Claim G.1. *To prove the lemma, it is sufficient to show that for every non zero constant polynomial $b \in \mathbb{F}_d[z]_k$, $\psi(b) = a$.*

The proof of Claim G.1 is given in Section G.1.1. Now we show that for every $b \in \mathbb{F}_d^\times, \psi(b) = a$ for some $a \in \mathbb{F}_d^\times$. Let B be a $(d-1) \times d$ size matrix, whose rows and columns are indexed by the ordered tuples $(\psi(1), \ldots, \psi(d-1))$ and $(0, \ldots, d-1)$ respectively and for $b \in \mathbb{F}_d^\times, \ell \in \mathbb{F}_d, B(b, \ell) := \psi(b)(\ell)$. Note that Observation G.1 implies that all the entries of B are non-zero.

Observation G.4. *To show $\psi(b) = a$ for every $b \in \mathbb{F}_d^\times$, it is sufficient to show that all the entries of B are a.*

Claim G.2. *There exists $a \in \mathbb{F}_d^\times$, such that every entry of B is equal to a.*

Proof. For $r \in [d] \setminus \{1,\dots,k\}$, let $h_r \in \mathbb{F}_d[z]_k$, such that $h_r(1) = \cdots = h_r(k-1) = h_r(r) = 0, h(k) = 1$. It is easy to see that for distinct $r_1, r_2 \in [d] \setminus \{1,\dots,k\}$, h_{r_1} and h_{r_2} are distinct polynomials.

Observation G.5. *For every $i \in \mathbb{F}_d \setminus \{1,\dots,k-1,r\}, h_r(i) \neq 0$. Further, for distinct $r_1, r_2 \in [d] \setminus \{1,\dots,k\}$ and for every $i \in \mathbb{F}_d \setminus \{1,\dots,k\}, h_{r_1}(i) \neq h_{r_2}(i)$.*

Fix $r \in [d] \setminus \{1,\dots,k\}$. Then Equation (17) implies that

$$\phi(h_r)(\sigma(1)) = \cdots = \phi(h_r)(\sigma(k-1)) = \psi(h_r)(\sigma(r)) = 0. \tag{19}$$

As $h_r(k) = 1$, Observations G.2 and G.3 imply

$$\phi(h_r)(\sigma(k)) = \phi(h_r(k))(\sigma(k)) = \phi(1)(\sigma(k)) = \psi(1)(\sigma(k)) = \psi(1)(\sigma(k)) \cdot h_r(k). \tag{20}$$

Set $a = \psi(1)(\sigma(k))$. The two equations above imply the following.

Observation G.6. *Let $r \in [d] \setminus \{1,\dots,k\}$. Then, for every $i \in \mathbb{F}_d, \phi(h_r)(\sigma(i)) = a \cdot h_r(i)$.*

Proof. As σ is linear, there exist $b_\sigma \in \mathbb{F}_d^\times, c_\sigma \in \mathbb{F}_d$, such that for every $i \in \mathbb{F}_d, \sigma(i) = b_\sigma \cdot i + c_\sigma$. Then, for every $i \in \mathbb{F}_d, \phi(h_r)(\sigma(i)) = \phi(h_r)(b_\sigma \cdot i + c_\sigma)$. Equations (19) and (20) imply that the polynomials $a \cdot h_r$ and $\phi(h_r)(b_\sigma \cdot z + c_\sigma)$ agree on $k+1$ points $1,\dots,k,r$. Since these are two degree atmost k polynomials, $a \cdot h_r = \phi(h_r)(b_\sigma \cdot z + c_\sigma)$. □

Fix $i \in \mathbb{F}_d \setminus \{1,\dots,k\}$. Observations G.2 and G.6 imply the following for every $r \in [d] \setminus \{1,\dots,k,i\}$ as $h_r(i) \neq 0$.

$$\phi(h_r)(\sigma(i)) = \sum_{l=1}^{h_r(i)} \psi(l)(\sigma(i)) = a \cdot h_r(i).$$

Let

$$S_{\sigma(i)} := \left\{ \sum_{l=1}^{h_r(i)} \psi(l)(\sigma(i)) = a \cdot h_r(i) \;:\; r \in [d] \setminus \{1,\dots,k,i\} \right\} \tag{21}$$

We claim that $S_{\sigma(i)}$ contains $d-k-1$ distinct equations. Suppose not, then there exist distinct $r_1, r_2 \in [d] \setminus \{1,\dots,k,i\}$ such that

$$\sum_{l=1}^{h_{r_1}(i)} \psi(l)(\sigma(i)) = \sum_{l=1}^{h_{r_2}(i)} \psi(l)(\sigma(i)).$$

This means $a \cdot (h_{r_1} - h_{r_2})(i) = 0$. As noted in Observation G.5, this happens if either $i \in \{1,\dots,k\}$ or $i = r_1 = r_2$, which is a contradiction. Using $S_{\sigma(i)}$, we show the following subclaim.

Subclaim G.1. *All the entries in the $\sigma(i)$-th column of B are a.*

The proof of Subclaim G.1 is given in Section G.1.1. This shows that for every $i \in \mathbb{F}_d \setminus \{1,\dots,k\}$, all the entries in the $\sigma(i)$-th column of B are a. Now we show that this is also the case with the remaining k columns $\sigma(1),\dots,\sigma(k)$. For $r \in [d] \setminus \{k+1,\dots,2k\}$ redefine h_r to be a polynomial in $\mathbb{F}_d[z]_k$, such that $h_r(r) = h_r(k+1) = \cdots = h_r(2k-1) = 0, h_r(2k) = 1$. Now, a subclaim similar to Subclaim G.1 shows that all the entries in some $d-k$ columns including the columns indexed by $\sigma(1),\dots,\sigma(k)$ are $\tilde{a}$ for some $\tilde{a} \in \mathbb{F}_d$. As $d \geq 4k+1$, there is atleast one column in B, where a and a' coincide. This implies that $a = \tilde{a}$. □

□

G.1.1 Proofs of a claim and a subclaim used in Lemma G.1

Proof of Claim G.1:
Let $h \in \mathbb{F}_d[z]_k$. We want to show that $\phi(h)$ satisfies Equation (18), i.e. for every $i \in \mathbb{F}_d$,

$$\phi(h)(\sigma(i)) = a \cdot h'(\sigma(i)) = a \cdot h(i) \qquad (p = 0 \text{ as } \phi(0) = 0)$$

Observe that if $\phi(h)(\sigma(i)) = 0$ then Equation (17) implies $h(i) = 0$. Suppose $\phi(h)(\sigma(i)) \neq 0$, then Observations G.2 and G.3, imply that

$$\phi(h)(\sigma(i)) = \phi(h(i))(\sigma(i)) = \sum_{l=1}^{h(i)} \psi(l)(\sigma(i)).$$

Since, $\psi(l) = a$ for every $l \in \mathbb{F}_d^\times$, we get $\phi(h)(\sigma(i)) = a \cdot h(i)$.

Proof of Subclaim G.1:

Proof. First we show that atleast $d - 2k - 1$ entries in the $\sigma(i)$-th column of B are a. Note that the $(d-k-1)$ equations in $S_{\sigma(i)}$ can be identified with $d-k-1$ distinct non-zero elements $\{h_r(i) : r \in [d] \setminus \{1, \ldots, k, i\}\}$. As i is fixed, for simplicity, we set $h_r(i)$ as b_r for $r \in [d] \setminus \{1, \ldots, k, i\}$. For some r, if the equation corresponding to b_r and $b_r + 1$ are in $S_{\sigma(i)}$ then observe that $\psi(b_r + 1)(\sigma(i)) = B(b_r + 1, \sigma(i)) = a$.

Suppose $r \in [d] \setminus \{1, \ldots, k, i\}$ be such that the equation corresponds to b_r is not in $S_{\sigma(i)}$ but the equations corresponding to $b_r - 1$ and $b_r + 1$ are in $S_{\sigma(i)}$. Then, we can not say if $\psi(b_r)(\sigma(i)), \psi(b_r + 1)(\sigma(i))$ are equal to a or not. Notice that there are k values of such r, such that corresponding b_r are non-zero. Thus, we do not know if $2k$ entries of the $\sigma(i)$-th column of B are a or not. Observe that for such distinct r_1, r_2, if b_{r_1}, b_{r_2} are consecutive elements in $\mathbb{F}_d^\times$, then the number of entries in the $\sigma(i)$-th column of B not known to be a is strictly less than $2k$. This shows that atleast $d - 2k - 1$ entries in the $\sigma(i)$-th column of B are a.

Now we show that the remaining entries in the $\sigma(i)$-th column are also a. Recall that $S_{\sigma(i)}$ contains equations corresponding to elements in $V := \{b_r : r \in [d] \setminus \{1, \ldots, k, i\}\}$. Pick $b' \in \mathbb{F}_d^\times \setminus V$ and $i' \in [d] \setminus \{1, \ldots, k, i\}$. Let $U \subseteq [d] \setminus \{i\}$, such that $i' \in U$, and $|U| = k$. For $s \in [d] \setminus U$, let $p_s \in \mathbb{F}_d[z]_k$ such that $p_s(i_1) = p_s(s) = 0$ for every $i_1 \in U \setminus \{i'\}$ and $p_s(i') = b'$. Then, from Equation (17), Observations G.2 and G.3, we get

$$\phi(p_s)(\sigma(i_1)) = \phi(p_s)(\sigma(s)) = 0 \text{ for every } i_1 \in U \setminus \{i'\} \text{ and}$$

$$\phi(p_s)(\sigma(i')) = \sum_{l=1}^{p_s(i')} \psi(l)(\sigma(i')).$$

Clearly, $\sum_{l=1}^{p_s(i')} \psi(l)(\sigma(i')) \neq 0$. Observe that for $i_1 \in U \cup \{s\}$,

$$\phi(p_s)(\sigma(i_1)) = a' \cdot p_s(i_1),$$

where $a' := \frac{\sum_{l=1}^{p_s(i')} \psi(l)(\sigma(i'))}{b'}$. It is immediate that $a' \neq 0$. An observation similar to Observation G.6 implies that for every $i_1 \in \mathbb{F}_d$,

$$\phi(p_s)(\sigma(i_1)) = a' \cdot p_s(i_1).$$

Then we get the following system of linear equations (similar to $S_{\sigma(i)}$ defined in Equation 21)

$$T_{\sigma(i),U} = \left\{ \sum_{l=1}^{p_s(i)} \psi(l)(\sigma(i)) = a' \cdot p_s(i) \ : \ s \in [d] \setminus U \right\}.$$

From the arguments used above, it follows that atleast $d - 2k - 1$ entries in the $\sigma(i)$-th column of B are a'. As $d - 2k - 1 > \frac{d}{2}$, we get $a = a'$. Observe that there exist different subsets U, such that $T_{\sigma(i),U} \neq S_{\sigma(i)}$ and $T_{\sigma(i),U} \cap S_{\sigma(i)}$ is also not empty. Thus, going over such subsets U, implies that all the entries in the $\sigma(i)$-th column of B are a. □

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il