A Dichotomy Theorem for Homomorphism Polynomials

N. de Rugy-Altherre

Univ Paris Diderot, Sorbonne Paris Cité,
Institut de Mathématiques de Jussieu, UMR 7586 CNRS,
F-75205 Paris, France
nderugy@math.univ-paris-diderot.fr

October 30, 2012

Abstract

In the present paper we show a dichotomy theorem for the complexity of polynomial evaluation. We associate to each graph H a polynomial that encodes all graphs of a fixed size homomorphic to H. We show that this family is computable by arithmetic circuits in constant depth if H has a loop or no edge and that it is hard otherwise (i.e., complete for VNP, the arithmetic class related to $\#P$). We also demonstrate the hardness over $\mathbb{Q}$ of cut eliminator, a polynomial defined by Bürgisser which is known to be neither VP nor VNP-complete in $\mathbb{F}_2$, if VP $\neq$ VNP (VP is the class of polynomials computable by arithmetic circuit of polynomial size).

1 Introduction

Dichotomy theorems are a way to classify entire classes of problems by their complexity. For instance, Schaefer [Sch78] has classified a subclass of CSP as being in P or NP-complete. Whether such a theorem holds for every CSP is a famous open question, the Dichotomy Conjecture [FV98]. More recently, dichotomy theorems for counting homomorphism problems have been studied by several authors [DG00, BG04, GT11]. In Valiant's theory of polynomial evaluation, an equivalent of Schaefer's theorem has been produced by Briquel and Koiran [BK09], opening the way for such theorems in this theory.

In this paper, we define the class of *homomorphism polynomials*: if H is a graph, the monomials of the associated homomorphism polynomial f_n^H encode the list of graphs of size n homomorphic to H. Our main result is a dichotomy theorem for this class of polynomials, in Valiant's theory : (f_n^H) is computable by unbounded fan-in arithmetic circuits in constant depth (the Valiant equivalent of AC_0) if H has a loop or no edge and hard otherwise (i.e., complete for VNP, the arithmetic class related to $\#P$).

Deciding the existence of or counting homomorphisms are important problems which generalize many natural questions (k-coloring, acyclicity, binary CSP, etc [HN04]). The weighted version, which is very similar to our polynomials, has several applications in statistical physics [Wel93].

The connection between Valiant's theory of polynomial evaluation and counting problems is well known. For instance, the permanent is complete in both settings [Val79a, Val79b]. Homomorphism polynomials are therefore linked to counting homomorphism problems: the evaluation of f_n^H counts the number of graphs of size n homomorphic to H. But these polynomials also tell us something about enumeration: it is equivalent to enumerate this set of graphs and to enumerate all the monomials of f_n^H. We will use this property to demonstrate a result on representations by first order formulas of the bipartite property for graphs.

The hard part of our dichotomy theorem is to demonstrate the hardness of most homomorphism polynomials. We therefore obtain many new VNP-complete families (indeed unlike NP-complete problems, there are not a lot of VNP-complete polynomials known). These families can be seen as generating functions in the way introduced by Jerrum and Bürgisser [Bür00]. No general hardness theorem has been found for these functions (but there are large results for tractable functions [CMR01]). As stated before, many problems can be seen as homomorphism problems. Our theorem therefore implies the hardness of many generating functions.

Moreover, Bürgisser [Bür00] gives a family of polynomials, Cut^2, which is neither VNP-complete nor VP over $\mathbb{F}_2$, if VNP $\neq$ VP of course. He also demonstrates the existence of such a family in every field, but does not give a specific family for $\mathbb{Q}$. He wonder whether Cut^2 can be such a family in $\mathbb{Q}$. During the demonstration of our theorem, we will show the VNP-completeness of Cut^2 and therefore answer negatively the question.

2 Definitions and general discussions

A polynomial enumerates a family $\mathcal{P}$ of graphs if each monomial of this polynomial encodes a single graph G in $\mathcal{P}$ and each graph $G \in \mathcal{P}$ is encoded by a single monomial. In this paper we study polynomials enumerating graphs G homomorphic to a graph H using a simple code: the variables of the polynomial are the set $\{x_e, e \in E(K_n)\}$; the monomial coding G is $\prod_{e \in E(G)} x^e$. These polynomials are therefore multilinear and have only 0 or 1 for coefficients. More precisely

Definition 1. Let H be a graph. The polynomial enumerating all graphs G with n vertices which are homomorphic to H is:

$$f_n^H((x_e)_{e \in E(K_n)}) := \sum_{\bar{\epsilon} \in \{0,1\}^{|E(K_n)|}} \Phi_n^H(\bar{\epsilon}) \prod_{e \in E(K_n)} x_e^{\epsilon_n}$$

Where $\Phi_n^H(\bar{\epsilon})$ is equal to 1 if the graph G such that $V(G) = [n]$ and $E(G) = \{e \in E(K_n) | \epsilon_e = 1\}$ is homomorphic to H, 0 otherwise.

We work within Valiant's algebraic framework. Here is a brief introduction to this complexity theory. For a more complete overview, see [Bür00].

An *arithmetic circuit* over a field $\mathbb{K}$ is a labeled directed acyclic connected graph with vertices of indegree 0 or 2 and only one sink. The vertices with indegree 0 are called *input gates* and are labeled with variables or constants from $\mathbb{K}$. The vertices with indegree 2 are called *computation gates* and are labeled with $\times$ or $+$. The sink of the circuit is called the *output gate*.

The polynomial computed by a gate of the arithmetic circuit is defined by induction : an input gate computes its label; a computation gate computes the product or the sum of its children's values. The polynomial computed by an arithmetic circuit is the polynomial computed by the sink of the circuit.

A *p-family* is a sequence (f_n) of polynomials over a field $\mathbb{K}$ such that the number of variables as well as the degree of f_n is polynomially bounded in n. The *complexity* $L(f)$ of a polynomial $f \in \mathbb{K}[x_1, \ldots, x_n]$ is the minimal number of computational gates of an arithmetic circuit computing f from variables $x_1, \ldots, x_n$ and constants in $\mathbb{K}$.

Two of the main classes in this theory are: the analog of P, VP, which contains of every p-family (f_n) such that $L(f_n)$ is a function polynomially bounded in n; and the analog of NP, VNP. A p-family (f_n) is in VNP iff there exists a VP family (g_n) such that for all n, $f_n(x_1, \ldots, x_n) = \sum_{\bar{\epsilon} \in \{0,1\}^n} g_n(x_1, \ldots, x_n, \epsilon_1, \ldots, \epsilon_n)$.

Furthermore we introduce the Valiant equivalent of AC_0, VAC_0 as the class of every p-family (f_n) such that there exists a circuit of constant depth but with unfounded fan-in for the computation gates computing f_n and of size polynomially bounded in n. It has been introduce in [MRar].

As in most complexity theory we have a natural notion of reduction: a polynomial f is a projection of a polynomial g, write $f \leq_p g$, if there are values $a_i \in \mathbb{K} \cup \{x_1, \ldots x_n\}$ such that $f(\bar{x}) = g(\bar{a})$. A p-family (f_n) is a p-projection of (g_n) if there exists a polynomially bounded function p such that for every n, $f_n \leq_p g_{p(n)}$. As usual we say that (g_n) is complete for a class C if, for every $(f_n) \in C$, $(f_n) \leq_p (g_n)$ and if $(g_n) \in C$.

The three classes presented above are closed under projections. We will need in this paper a second notion of reduction, the c-reduction: the oracle complexity $L^g(f)$ of a polynomial f with the knowledge of g is the minimum number of computation gates and evaluations of g over previously computed values that are sufficient to compute f from the variables $x_1, \ldots x_n$ and constants from $\mathbb{K}$. A p-family (f_n) *c-reduces* to (g_n) if there exists a polynomially bounded function p such that $L^{g_{p(n)}}(f_n)$ is a polynomially bounded function.

This reduction is more powerful than the projection: if $(f_n) \leq_p (g_n)$, then $(f_n) \leq_c (g_n)$. With this reduction, VNP is still closed, but it is harder to demonstrate (See [Poi08] for a idea of the proof). However this reduction does not distinguish the lower classes. For example, 0 is VP-complete for c-reductions.

The notion of c-reductions was introduced by Bürgisser [Bür00] in his work on p-families that are neither VP nor VNP-complete (if VNP $\neq$ VP). It has been rarely used to demonstrate the hardness of computing polynomials, though there is a recent example in [Bri11].

Bürgisser demonstrates the existence of such a family in every field with an abstract embedding theorem. Furthermore he gives a specific family for some finite field, the *cut enumerator* Cut_n^q:

$$\text{Cut}_n^q := \sum_S \prod_{i \in A, j \in B} x_{ij}^{q-1}$$

where the sum is over all cuts $S = \{A, B\}$ of the complete graph K_n on the set of nodes $\{1, \ldots, n\}$. (A cut of a graph is a partition of its set of nodes into two nonempty subsets).

Theorem 1. *(Bürgisser, 5.22) Let q be a power of the prime p. The family of cut enumerators Cut^q over a finite field $\mathbb{F}_q$ is neither VP nor VNP-complete with respect to c-reductions, provided $Mod_p\text{NP} \nsubseteq \text{P}/poly$. The latter condition is satisfied if the polynomial hierarchy does not collapse at the second level.*

Howerver, the cut enumerator can be rewritten in the following way :

$$\mathrm{Cut}_n^q(\bar{x}) = \sum_{\substack{V \subsetneq [n] \\ V \neq \emptyset}} \prod_{i \in V} \prod_{j \in V^c} x_{i,j}^{q-1}$$

This polynomial will be studied in Lemma 7, for other purposes. In particular, we will demonstrate that this family is VNP-complete in $\mathbb{Q}$ if $q = 2$, which answers a question of Bürgisser. Furthermore, this implies that our demonstration does not work in a field of characteristic 2 (if VP $\neq$ VNP).

The main theorem of our paper is:

Theorem 2. *Let H be a graph. Let f_n^H be the polynomial enumerating all graphs homomorphic to H. Then, over $\mathbb{Q}$*

- *If H has a loop or no edges, (f_n^H) is in VAC_0.*

- *Else (f_n^H) is VNP-complete for c-reductions.*

Proofs that all our p-families are p-families and are in VNP are left to the reader. It is a simple application of Valiant's criterion [Bür00]. We will demonstrate our main theorem in four steps: first we will deal with the easy cases. Secondly, using Dyer and Greenhill's ideas, we will reduce the problem to the case where H is as simple as possible, i.e., with only one edge and two vertices, •—•. In a third step, we will note that the monomial of $f^{•—•}$ have a kind of "hereditar" property which we will use to reduce the problem to a simpler polynomial. Finally, we will demonstrate the VNP-completeness of this simpler polynomial.

As application of our theorem and using result on enumeration, we can prove that:

Corollary 1. *If $\mathrm{VP} \neq \mathrm{VNP}$, and if the product of two boolean matrices of size n cannot be computed in $\mathcal{O}(n^2)$, being bipartite cannot be expressed by an acyclic conjunctive first order formula of polynomial size on the size of the studied graph.*

We just give here an idea of the demonstration. Let ϕ_n be an acylic conjunctive first order formula of polynomial size stating that G is bipartite or not, for every graph G of size n. As the set of bipartite graphs can be enumerated with a constant delay, by the dichotomy theorem of Bagan [Bag09], ϕ_n is CCQ, i.e., of star size 1. Therefore, thanks to the work of Durand and Mengel [DM11], we know that $P(\phi) \in \mathrm{VP}$. At last, as $P(\phi) = f^{•—•}$, $f^{•—•}$ is VNP-complete and in VP, which is impossible if $\mathrm{VNP} \neq \mathrm{VP}$. Therefore ϕ_n cannot exist.

3 The easy cases

Lemma 1. *If H has a loop, then (f_n^H) is in VAC_0.*

Proof. Let v be a looped vertex of H. Then every graph G is homomorphic to H: we just have to send all vertices of G to v. Therefore,

$$f_n^H(\bar{x}) = \sum_{\bar{\epsilon} \in \{0,1\}^{|E(K_n)|}} \prod_{e \in E(K_n)} x_e^{\epsilon_n} = \prod_{e \in E(K_n)} (1 + x_e)$$

This product can be computed easily by an arithmetic circuit of polynomial size, depth 3 and unbounded fan-in for the $\times$-gates. □

The case where H has no edges is quite obvious, we just have to notice that if H_1 and H_2 are two bihomomorphic graphs, i.e., there exists a homomorphism from H_1 to H_2 and vice versa, then, for all n, $f_n^{H_1} = f_n^{H_2}$. Let us consider now our graph H with no edges. It is bihomomorphic to the graph H_0, composed of a single vertex. Furthermore the only graphs homomorphic to H_0 are those with no edges. Therefore $f_n^{H_0} = 1$.

Let us quote a very useful technical result, already known in [Bür00].

Lemma 2. *Let (f_n) be a p-family. Let us write $\mathbf{HC}_k(f_n)$ for the homogeneous component of degree k of f_n.*

Then for any sequence of integers (k_n) there exists a c-reduction from the homogeneous component to the polynomial itself:

$$(\mathbf{HC}_{k_n}(f_n)) \leq_c (f_n)$$

Proof. Let n be an integer and d_n be the degree of f_n. Let us recall that the degree of f_n is polynomially bounded in n. To begin, let us notice that for all i, $f_n(2^i \bar{x}) = \sum_{k=0}^{d_n} (2^i)^k \mathbf{HC}_k(f_n)(\bar{x})$. If we write $c_k = \mathbf{HC}_k(f_n)(\bar{x})$, $f_i = f_n(2^i \bar{x})$ and $w_k = 2^k$, those equations can be written as:

$$
\begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_{d_n} \end{pmatrix} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \omega_1 & \omega_2 & \cdots & \omega_d \\ \vdots & \vdots & \vdots & \vdots \\ \omega_1^{d_n} & \omega_2^{d} & \cdots & \omega_{d_n}^{d_n} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_{d_n} \end{pmatrix}
$$

Let us write V the square matrix of this equation. It is a Vandermonde matrix with non negative integers. As the coefficients w_k are distinct, this matrix is invertible. Therefore there exist some rationals $w_{i,k}^*$ such that $c_k = \sum_{i=0}^{d_n} w_{i,k}^* f^n(2^i \bar{x})$.

So we have, for all n and all $k \le d_n$, $L^{f_n}(\mathbf{HC}_k(f_n)) = \mathcal{O}(n \times d_n)$. Let (k_n) be a sequence of integers. If $k_n > d_n$, then $\mathbf{HC}_{k_n}(f_n) = 0$. Thus for all $L^{f_n}(\mathbf{HC}_{k_n}(f_n)) \le \mathcal{O}(n \times d_n)$. As d_n is polynomially bounded by n, we have just constructed the c-reduction we were looking for. $\qquad\square$

This lemma is still true if we take the homogeneous compenent in some variables only. Let $f(\bar{x}, \bar{y})$ be a polynomial and $\mathbf{HC}_k^{\bar{y}}(f)$ the homogeneous compenent of size k in variables $\bar{y}$, i.e., considering $\bar{x}$ as constant. Then for any p-family (f_n) and any sequence of integers k_n, $\left(\mathbf{HC}_k^{\bar{y}}(f_n)\right) \le_c (f_n)$.

Remark. *If C is a circuit that compute of size s a polynomial f of degree d, we know that there exists an arithmetic circuit of size $d^2 s$ that compute all homogeneous component of f (Theorem 2.2 in [SY10] for example). With the previous lemma we can see that if C is a arithmetic formula of size s that compute f of degree d, then for any $k \le d$, we can built an arithmetic formula of size $\mathcal{O}(d^2 s)$ that compute the homogeneous component of degree k of f (this formula must use constant of $\mathbb{Q}$).*

4 Reduction to bipartite graphs

One of the main idea of Dyer and Greenhill is to reduce the problem of counting homomorphisms from G to H to a simpler graph H. Here we have a similar result but the graph we obtain is extremely simple.

Proposition 1. *Let $\bullet\!\!-\!\!\bullet$ be the graph with two vertices and one edge. Then for any graph H with no loops and at least one edge:*

$$f_n^{\bullet\!-\!\bullet} \le_c f_n^H$$

Lemma 3. *Let $H = (V, E)$ be a graph with no loops and at least one edge. For all $i \in V$, let us call H_i the subgraph of H induced by the neighbours of i and H' the disjoint union of all those neighbourhood graphs. Then*

$$f_n^{H'} \le_c f_n^H$$

Proof. For any graph G, let us call G' the graph built from G by adding a new vertex v and joining it to every vertex of G. Dyer and Greenhill have noticed in their paper [DG00] that the number of homomorphisms from G' to H is equal to the number of homomorphisms from G to H'. Let us see what this means for our polynomials.

Let $\tilde{f}_n^H(\bar{x})$ be the polynomial built from f_{n+1}^H by replacing all $x_{i,n+1}$ variables by y, by taking the homogeneous component of degree n in y and by giving value 1 to y: $\tilde{f}_n^H(\bar{x}) = \mathbf{HC}_n^y \left(f_{n+1}^H(\bar{x}, y) \right)_{|y=1}$.

$\tilde{f}_n^H$ enumerates all graphs of size $n + 1$ homomorphic to H and having a vertex v linked to all others, i.e., graphs G' homomorphic to H.

$$\tilde{f}_n^H(\bar{x}) = \sum_{G \text{ graph of size } n} \Phi_n^H(G') \bar{x}^{\bar{G}} = \sum_{G \text{ graph of size } n} \Phi_n^{H'}(G) \bar{x}^{\bar{G}} := f_n^{H'}(\bar{x})$$

where $\Phi_n^H(G) = 1$ if there exists a homomorphism between G and H, 0 otherwise. Then $\tilde{f}_n^H = f_n^{H'}$ and consequently there is a c-reduction from $(f_n^{H'})$ to (f_{n+1}^H). $\qquad\square$

Proof of the proposition 1. Since H has no loops, the degree of H' is strictly lower than that of H. Indeed, for any vertex j of H, any vertex k linked to j in H' is also linked to j in H. So the degree of H' cannot grow. Furthermore, if i is a maximal degree (d) vertex of H and if j is a vertex of H, then either i is not linked to j and so does not appear in H_j, or it is linked to j and, in H_j, it cannot be linked to more than $d - 1$ vertices ($|V_i \cap V_j|$, knowing that $j \notin V_j$). The degree of the vertices of H' corresponding to the vertex i is therefore strictly lower than d.

If we apply the process of Lemma 3 to H a finite number of times we obtain a 1-regular graph. This graph is bihomomorphic to $\bullet\!\!-\!\!\bullet$. Therefore we have the expected reduction: $f_n^{\bullet\!-\!\bullet} \le_c f_n^H$ $\qquad\square$

5 The bipartite case

Now we have to demonstrate the following proposition.

Proposition 2. *Let* $\bullet\!\!-\!\!\bullet$ *be the graph with two vertices and one edge. Then* $f_n^{\bullet\text{-}\bullet}$ *is VNP-complete.*

5.1 Hereditary polynomials

$f_n^{\bullet\text{-}\bullet}$ has a noteworthy property: if $(x_i)_{i\in I}$ is a monomial of $f_n^{\bullet\text{-}\bullet}$, then for all $J \subset I$, $(x_j)_{j\in J}$ is also a monomial of $f_n^{\bullet\text{-}\bullet}$. Indeed, this polynomial enumerates all graphs homomorphic to $\bullet\!\!-\!\!\bullet$, i.e., all bipartite graphs, and every subgraph of a graph homomorphic to $\bullet\!\!-\!\!\bullet$ is also homomorphic to $\bullet\!\!-\!\!\bullet$. To build on this idea, we introduce the following definitions.

Let f be a multilinear polynomial with 0,1 coefficients only. f is *hereditary* if it satisfies the following property: if $(x_i)_{i\in I}$ is a monomial of f, so is $(x_j)_{j\in J}$, for all $J \subset I$. This relation induces an order on all the monomials of f: $(x_i)_{i\in I}$ is greater that $(x_j)_{j\in J}$ if $J \subset I$. We call a monomial a *generator* for f if it is maximal for this order.

A hereditary polynomial is completely defined by its generators. We can naturally ask the opposite question. Let g be a multilinear polynomial with 0,1 coefficients. We write $\downarrow g$ for the *son* of g, i.e., the smallest hereditary polynomial containing g.

Lemma 4. *Let* g_n *be a sequence of multilinear homogeneous polynomials with* 0,1 *coefficients. Then*

$$(g_n) \leq_c (\downarrow g_n)$$

Proof. Let n be a integer and d the degree of all the monomials of g_n. The generators of $\downarrow g_n$ are simply the monomials of degree d, $g_n = \mathbf{HC}_d(\downarrow g_n)$. The reduction is given by Lemma 2. $\square$

The reduction $(g_n) \leq_c (\downarrow g_n)$ is probably not true in the general case: for instance $f_n = (\prod_{i=1}^n x_i) + per(\bar{x})$ is VNP-complete, but its son is the sum of all monomials of size less than n, and it is therefore in VAC_0. However, we can hope to generalize this equality to all *pure* polynomials, which we define as polynomials with no pair of comparable monomials.

In our case, we try to demonstrate that $f_n^{\bullet\text{-}\bullet}$ is VNP-complete. This polynomial is hereditary. In Lemma 6, we will demonstrate that the polynomial F_n, which enumerates all graphs composed of some isolated vertices and a complete bipartite connected component, is VNP-complete. Now $f_n^{\bullet\text{-}\bullet}$ is the son of this polynomial: $\downarrow F_n = f_n^{\bullet\text{-}\bullet}$.

Unfortunately this polynomial F_n is not pure, some of its monomials are comparable. In Lemma 8 we demonstrate that the polynomial (G_n), enumerating all the complete bipartite graphs, is also VNP-complete. $f_n^{\bullet\text{-}\bullet}$ is also the son of G_n, but this polynomial is pure!

As we have no general proof of the reduction from a pure polynomial to its son (yet), we will now give an ad hoc demonstration of the reduction from G_n to $f_n^{\bullet\text{-}\bullet}$.

At the end, we will have demonstrated the following chain, for any graph H with no loop and at least one edge.

$$(\text{GF}(K_n, \text{clique})) \leq_c (F_n) \leq_p (\text{Cut}_n^2) \leq_c (G_n) \leq_c (f_n^{\bullet\text{-}\bullet}) \leq_c (f_n^H)$$

As the generating function of clique is VNP complete, we will have demonstrated Theorem 2

Lemma 5. *Let* $f_n^{\bullet\text{-}\bullet}$ *be the polynomial enumerating all graphs homomorphic to* $\bullet\!\!-\!\!\bullet$. *Let* G_n *be the polynomial enumerating all complete bipartite graphs. Then*

$$G_n \leq f_n^{\bullet\text{-}\bullet}$$

Proof. G_n can be written as $G_n\left((y_e)_{e\in E(K_n)}\right) = \frac{1}{2}\sum_{V\subset[n]}\prod_{v\in V}\prod_{v'\in V^c} y_{v,v'}$. We express G_n in the following way:

$$G_n = \left(\sum_{k=1}^n \mathbf{HC}_{(n-k)k}^{\bar{x}}\left(\mathbf{HC}_k^y\left(\mathbf{HC}_{n-k}^z\left(\mathbf{HC}_1^w\left(f_{\bullet\text{-}\bullet}^{n+2}(\bar{x})\right)\right)\right)\right)\right)_{|w,y,z=1}$$

Where $x_{n+1,n+2} = w$ and for all $i \in [n]$, $x_{i,n+1} = y$ and $x_{i,n+2} = z$.

Indeed, G_n enumerates all bipartite graphs. Each operation above consisting in taking a homogeneous component eliminates some graphs from the list. We will demonstrate that, at the end, the polynomial will only enumerate complete bipartite graphs.

In a first step, we add two vertices $n+1$ and $n+2$ and we give a weight w to the edge between them. By taking the homogeneous component of degree 1 in the variable w, we eliminate all graphs but those which have an edge between $n+1$ and $n+2$.

Then we use variable substitutions to label the edges leaving $n+1$ with the weight y and those leaving $n+2$ with z. A vertex linked to $n+1$ cannot be linked to $n+2$, as our graphs are bipartite and the vertices $n+1$ and $n+2$ are linked. Let us write V and V' for the partition of vertices defined by the bipartite graph, with $n+2 \in V$ and therefore, as $n+1$ and $n+2$ are linked, $n+1 \in V'$

By taking next the homogeneous component of degree $n - k$ in z, we force that $|V'| \geq n - k$. Similarly, by taking the homogeneous component of degree k in y, we force that $|V| \geq k$. Therefore, as V and V' are disjoint sets, we force that $V^c = V'$.

At the end, by taking the homogeneous component of degree $(n - k)k$ in the variables $\bar{x}$, we only keep bipartite graphs built on V and V^c, with $|V| = k$ and which have $(n - k)k$ edges, i.e., complete bipartite graphs on n vertices.

The reduction is then given by Lemma 2. $\qquad \square$

5.2 Bipartite graphs

Let us demonstrate that F_n is VNP-complete by reducing the generating function of clique to F_n. Recall that $\downarrow F_n = f_n^{\bullet \bullet}$ but that F_n is not pure.

Lemma 6. *Consider the polynomial*

$$F_n \left((x_e)_{e \in E(K_n)} \right) = GF(K_n, \text{complete bipartite}) = \sum_{E'} \prod_{e \in E'} x_e,$$

where the sum is done over all subsets $E' \subset E(K_n)$ such that the graph (n, E') (i.e., the graph with n vertices and E' as set of edges) has one complete bipartite connected component and all other connected components reduce to a single vertex.

This p-family is VNP-complete for c-reductions.

Proof. We will reduce the clique generating function to our polynomial: $g_n(\bar{x}) = GF(K_n, \text{clique})((x_e)_{e \in E(K_n)}) = \sum_{E'} \prod_{e \in E'} x_e$,, where the summation is on all $E' \subset E(K_n)$ such that the graph (n, E') has a complete connected component (a clique) and all others reduce to a single vertex. Bürgisser has demonstrated in [Bür00] that this sequence is VNP-complete for p-reductions.

Let us rewrite our polynomial:

$$F_n(\bar{x}) = \frac{1}{2} \sum_{\substack{V, V' \subset [n] \\ V \cap V' = \emptyset}} \prod_{v \in V} \prod_{v' \in V'} x_{(v, v')}$$

The $\frac{1}{2}$ comes from the fact that each complete bipartite graph can be defined by two different couples of sets: (V, V') and (V', V); and that each couple of vertex sets correspond to a unique complete bipartite graph.

We will proceed to the reduction by beginning with the sequence (F_n) and writing a series of sequences of polynomials $(f_n^k)_{n \in \mathbb{N}}$, each being reducible to the previous one.

Hereafter, we suppose the left vertices of $K_{n,n}$ being $[1, n]$ and the right ones being $[n + 1, 2n]$. We write $\hat{i}$ for the i-th right vertex, i.e., $\hat{i} = i + n$. Similarly, if $[n]$ are the left vertices of $K_{n,n}$, we write $\widehat{[n]}$ for the right ones, i.e., $\widehat{[n]} = [n + 1, 2n]$.

$$f_n^1(\bar{x}) = F_{2n}(K_{n,n}) = \sum_{V \subset [n]} \sum_{V' \subset \widehat{[n]}} \prod_{i \in V} \prod_{j \in V'} x_{i,j}$$

Let V and V' be two disjoint sets of vertices. If V has some of its vertices both in the right side and the left side (let us say for instance i in the left side), then we can suppose that V' has a vertex j in the left side. Hence $x_{(i,j)} = 0$ appears in the weight of V', V. Therefore this weight is zero.

The only complete bipartite graphs enumerated in $F_{2n}(K_{n,n})$ are those built on V and V' with V a subset of the left vertices and V' a subset of the right ones. The summation on all the V subset of the left vertices avoids repetitions and thus the $\frac{1}{2}$ coefficient is not necessary.

$$f_n^2(\bar{x}, y) = f_{n+1}^1(\bar{x}, x_{n+1, \hat{i}} = 0, x_{i, \widehat{n+1}} = y) = \sum_{V \subset [n+1]} \sum_{V' \subset \widehat{[n+1]}} \prod_{i \in V} \prod_{j \in V'} x_{i,j}$$

If V and V' are vertex subsets, let $\omega(V, V')$ be their weight: $\omega(V, V') := \prod_{i \in V} \prod_{j \in V'} x_{i,j}$. Let $V \subset [n]$.

- If $n + 1 \in V$ then $\forall V' \subset \widehat{[n+1]}$, $\omega(V, V') = 0$.
- If $\widehat{n+1} \in V'$ then $\forall V \subset [n]$, $\omega(V, V') = (\prod_{i \in V} \prod_{j \in V' - \{\widehat{n+1}\}} x_{i,j}) \times y^{|V|}$.
- Else, for all $V \subset [n]$ $\deg_y(\omega(V, V')) = 0$.

In those polynomials, the degree of y is a witness to the size of V. We now just have to take the homogeneous component in y of degree k to keep all the V of size k:

$$f_n^{3,k}(\bar{x}) = \mathbf{HC}_k^y \left(f_n^2 \right) (\bar{x}, y = 1) = \sum_{\substack{V \subset [n] \\ |V| = k}} \sum_{V' \subset \widehat{[n]}} \omega(V, V')$$

Let us act similarly for V'.

$$f_n^{4,k}(\bar{x}, y) = f_{n+1}^{3,k}(\bar{x}, x_{\widehat{n+1,i}} = y, x_{i,\widehat{n+1}} = 0)$$

Here the degree of y is a witness to the size of V' in monomials where y is of degree at least 1. Then

$$f_n^{5,k}(\bar{x}) = \mathbf{HC}_k^y\left(f_n^{4,k}\right)(\bar{x}, 1) = \sum_{V \subset [n], |V| = k} \sum_{V' \subset \widehat{[n]}, |V'| = k} \omega(V, V')$$

Let us write now

$$f_n^{6,k}(\bar{x}, y) = f_n^{5,k}(\bar{x}, x_{i,i'} = y) = \sum_{V \subset [n], |V| = k} \sum_{V' \subset \widehat{[n]}, |V'| = k} \prod_{i \in V} \prod_{\substack{j \in V' \\ j \neq i}} x_{i,j} \times y^{|V \cap V'|}$$

And at last, we find the clique generating function:

$$g_n(\bar{x}) = \sum_{k=1}^n \mathbf{HC}_k^y\left(f_n^{6,k}\right)(\bar{x}, 1)$$

Indeed, in this polynomial we sum on all subsets $V \subset [n]$ of size k and the degree of y is a witness to the size of $V \cap V'$. Then if we keep only monomials for which the degree in y is k, we only keep complete bipartite graphs built on V and V', with $|V| = |V'| = |V \cap V'|$, i.e., on the same subset of vertices V, on the left and on the right. Then:

$$\sum_{k+1}^n \mathbf{HC}_k^y\left(f_n^{6,k}\right)(\bar{x}, 1) = \sum_{V \subset [n]} \prod_{i \in V} \prod_{j \in V} x_{i,\widehat{j}}$$

This polynomial is, by definition, $g_n(\bar{x})$, the clique generating function. Furthermore, as we announce it at the beginning of this demonstration, we have the following reductions:

$$f_n^2 \leq_p f_n^1 \leq_p F_n$$

And, for all n, $L^{f_{n+1}^2}(f_n^{3,k}) = \mathcal{O}(nk)$ then $L^{f_{n+1}^2}(f_n^{4,k}) = \mathcal{O}(nk)$. Similarly $L^{f_{n+1}^{4,k}}(f_n^{5,k}) = \mathcal{O}(nk)$ then $L^{f_{n+2}^2}(f_n^{6,k}) = \mathcal{O}(n^2 k^2)$. At last $L^{f_{n+2}^2}(g_n) = \sum_{k=1}^n \mathcal{O}(n^3 k^3) = \mathcal{O}(n^7)$.

Thus we have shown a c-reduction from g_n to F_n. $\qquad\square$

As we mentioned before the son of F_n is $f_n^{\bullet\bullet}$ (i.e., $\downarrow F_n = f_n^{\bullet\bullet}$), but F_n is not pure. We therefore had to introduce the polynomial G_n which is still a generator of $f_n^{\bullet\bullet}$ but is pure. For now we introduce another polynomial, the cut enumerator. We will use it to prove the completeness of G_n. But as mentioned in the introduction, theorem 1, this family had been introduced by Bürgisser [Bür00] as a family neither VP nor VNP-complete in $\mathbb{F}_2$, if of course VP $\neq$ VNP.

Be aware that, contrary to the rest of this paper, this polynomial will be built on non symmetric variables: $x_{i,j}$ will, generally, not be equal to $x_{j,i}$.

Lemma 7. *Let Cut_n^2 be the following polynomial*

$$Cut_n^2\left((x_{i,j})_{i,j \in [n]}\right) = \sum_{V \subset [n]} \prod_{v \in V} \prod_{v' \in V^c} x_{v,v'}$$

This p-family is VNP-complete for c-reductions.

Proof. We will exhibit a p-reduction from $F_n = \mathrm{GF}(K_n, \text{complete bipartite})$ to Cut_n^2. Because $\mathrm{GF}(K_n, \text{complete bipartite})$ is complete for c-reductions, as seen in Lemma 6, this will prove the lemma.

Let us suppose that the left vertices of $K_{n,n}$ are $[n]$, the right ones $[n+1, 2n]$ and let us write $\widehat{i}$ for the i-th right vertex, i.e., $\widehat{i} = i + n$. Similarly, if $[n]$ is the set of left vertices, we will write $\widehat{[n]}$ for the right ones. Let $(y_{i,j})_{i,j \in [2n]}$ be a set of new variables. We now give new values to the variables $x_{i,j}$ for all $i, j \in [n]$:

1. $\forall i, j \in [n] \; x_{i,j} = 1$

2. $\forall i \in [n] \; x_{i,\widehat{i}} = 0$

3. $\forall i \in [n] \; x_{\widehat{i,i}} = 1$

4. $\forall i \in [n] \forall j \in \widehat{[n]} \; x_{j,i} = 1$

5. $\forall i \in \widehat{[n]} \forall j \in \widehat{[n]} \; x_{i,j} = 1$

6. $\forall i \in [n] \forall j \in \widehat{[n]} \; x_{i,j} = y_{i,j}$

7

If we evaluate Cut^2_{2n} on those variables, we compute $\mathrm{GF}(K_n, \text{complete bipartite})$. Indeed, let $V \subset V(K_{n,n})$

- If there exists $i \in V$ such that $\widehat{i} \notin V$, then $\widehat{i} \in V^c$ and as $x_{i,\widehat{i}} = 0$ by (2), then $\omega(V) := \prod_{v \in V} \prod_{v' \in V^c} x_{v,v'} = 0$.

- If there exists $i \notin V$ such that $\widehat{i} \in V$, then $i \in V^c$ and as $x_{\widehat{i},i} = 1$ by (3), the value of $\omega(V)$ is not necessarily zero.

V is therefore of type $V_1 \cup V_2$ with V_1 in the left side, V_2 in the right one and $V_1' := \{\widehat{n}, n \in V_1\} \subset V_2$. Then if we write V_1^c the complement of V_1 in the left side and V_2^c the one of V_2 in the left side,

$$
\begin{aligned}
\omega(V) &= \prod_{v \in V} \prod_{v' \in V^c} x_{v,v'} \\
&= (\prod_{i \in V_1} \prod_{v' \in V^c} x_{i,v'})(\prod_{i' \in V_2} \prod_{v' \in V^c} x_{i',v'}) \\
&= (\prod_{i \in V_1} \prod_{j \in V_1^c} x_{i,j})(\prod_{i \in V_1} \prod_{j' \in V_2^c} x_{i,j'})(\prod_{i' \in V_2} \prod_{v' \in V^c} x_{i',v'})
\end{aligned}
$$

In the last line, the first parenthesis, $(\prod_{i \in V_1} \prod_{j \in V_1^c} x_{i,j})$, takes value 1 because of (1). The second one, $(\prod_{i \in V_1} \prod_{j \in V_2^c} x_{i,j})$, takes value $(\prod_{i \in V_1} \prod_{j \in V_2^c} y_{i,j})$ by (6).The last one, $(\prod_{i \in V_2} \prod_{v' \in V^c} x_{i,v'})$, take value 1 because of (4) and (5).

Then $\omega(V) = \prod_{i \in V_1} \prod_{j \in V_2^c} y_{i,j}$. But there is a bijection between $\{V \subset V(K_{n,n}), \omega(V) \neq 0\}$ and $\{(V, V'), V \subset [n], V' \subset [n].V \cap V' = \emptyset\}$. Furthermore this bijection conserves weights. Thus, for the values of the variables $\bar{x}$ defined above,

$$
\mathrm{Cut}^2_{2n}(\bar{x}) = \mathrm{GF}(K_n, \text{complete bipartite})(\bar{y}) = F_n
$$

$\square$

Lemma 8. *Let G_n be the polynomial enumerating all complete bipartite graphs of size less than n. Then the p-family (G_n) is VNP-complete for c-reductions.*

Proof. Classically, to pass from non oriented graphs to oriented graphs we double the number of vertices, K_{2n}, we call the n-th first as left vertices and the n-th last as right vertices. Then we labeled the edge between the i-th left vertex and the j-th right's by the oriented variable $x_{i,j}$. The edges between vertices of the same side are labeled with 1 and horizontal edges, i.e., edges between the i-th left vertex and the i-th right's are labeled with 0, as we do not want loops. Let us write K' this new graph.

G_n evaluate on this graph will enumerate every complete bipartite graphs. By taking homogeneous component we will reduce this list until we find the cut enumerator.

First, let us add to K' two new vertices, a and b. The new edge we create are labeled by:

- The edge between a and b is labeled by a new variable t,

- The edge between a and any left vertex by y,

- The edge between a and any right vertex by y',

- The edge between b and any left vertex by z,

- The edge between b and any right vertex by z',

By taking the homogeneous component of size 1 in t, we only keep in our list the graphs which have a edge between a and b.

For a $k \leq n$, we slip the left side of the graph K' into two sets: V and V^c by joining k vertices of the left side to a and $n - k$ left vertices to b (i.e., by taking the homogeneous component of size k in y and the homogeneous component of size $n - k$ in z). A vertex cannot be linked to both a and b, as our graphs must always be bipartite. We name V the set of left vertices linked to a, and therefore V^c is the set of left vertices linked to b.

Similarly, we split the right side of the graph K' into two sets: V', the set of the k right vertices linked to a and V'^c the set of the $n - k$ right vertices linked to b.

A vertex in V cannot be linked to a vertex in V' as they are both linked to a. Similarly a vertex in V^C cannot be linked to a vertex in V'^C because they are both linked to b. Furthermore, every edge between a vertex of V and a vertex of V^c, or between a vertex of V' and a vertex of V'^c is labeled by 1, as they are on the same side. Therefore, the only edge labeled with a $x_{i,j}$ variable that participate in G_n are those between a vertex of V and a vertex of V'^c and those between one of V' and one of V'^c.

If the i-th left vertex of K' is in V and the i-th right vertex of K' is in V'^c, as G_n enumerate only complete bipartite graphs, the monomial which represent this graph will have $x_{i,i} = 0$ and therefore will be null. Thus if we identify both the left and the right side to $[n]$, $V = V'$. Finally, if we evaluate G_n on K', and if evaluate the variables t, y, y', z, and z' to 1, we have

$$\sum_{V \subset [n]} \left(\prod_{i \in V} \prod_{j \in V'^c} x_{i,j} + \prod_{i \in V'} \prod_{j \in V^c} x_{i,j} \right) = 2\mathrm{Cut}^2(\bar{x})$$

□

Perspectives

We have left open in section 5.1 the question of whether a pure multilinear polynomial is always reducible to its son, though we have shown the reduction for homogeneous polynomials and in the special case of homomorphism polynomials. This question can be seen as a generalization of the relationship between the permanent and the partial permanent PER* (cf [Bür00]). Note that the partial permanent is useful to show the completeness of other families of polynomials, so it would be interesting to obtain a general result.

The opposite question, namely whether the son of a polynomial always reduces to it, is also quite interesting. For instance, PER is in VP in characteristic 2. The same result for PER* was implicitly solved in [Val01]. A reduction from $\downarrow f$ to f in the general case would provide another proof of this result and extend our understanding of these polynomials.

Another question raised by this paper is the relation between polynomials which list a graph property as homomorphic polynomials or generating functions on one hand and enumeration on the other. A brief foray into this subject has been made with corollary 1.

We intend to extend the study of this two questions in ours future researches.

Acknowledgements

References

[Bag09] Guillaume Bagan. *Algorithmes et complexité des problèmes d'énumération pour l'évalution de requêtes logique.* PhD thesis, Université de Caen/Basse-Normandie, 2009.

[BG04] Andrei Bulatov and Martin Grohe. The complexity of partition functions. In *Automata, Languages and Programming*, volume 3142 of *Lecture Notes in Computer Science*, pages 294–306. Springer Berlin / Heidelberg, 2004. 10.1007/978-3-540-27836-8-27.

[BK09] Irénée Briquel and Pascal Koiran. A dichotomy theorem for polynomial evaluation. In *MFCS*, pages 187–198, 2009.

[Bri11] Irénée Briquel. *Complexity issues in counting, polynomial evaluation and zero finding.* PhD thesis, Lyon, 2011.

[Bür00] Peter Bürgisser. *Completeness and Reduction in Algebraic Complexity Theory.* Springer, 2000.

[CMR01] B. Courcelle, J.A. Makowsky, and U. Rotics. On the fixed parameter complexity of graph enumeration problems definable in monadic second-order logic. *Discrete Applied Mathematics*, 108(1–2):23 − 52, 2001.

[DG00] Martin E. Dyer and Catherine S. Greenhill. The complexity of counting graph homomorphisms. *Random Struct. Algorithms*, 17(3-4):260–289, 2000.

[DM11] Arnaud Durand and Stefan Mengel. On polynomials defined by acyclic conjunctive queries and weighted counting problems. *CoRR*, abs/1110.4201, 2011.

[FV98] Tomas Feder and Moshe Y. Vardi. The computational structure of monotone monadic snp and constraint satisfaction: A study through datalog and group theory. *SIAM Journal on Computing*, 28(1):57–104, 1998.

[GT11] Martin Grohe and Marc Thurley. Counting homomorphisms and partition functions. *CoRR*, abs/1104.0185, 2011.

[HN04] Pavol Hell and Jaroslav Nesetril. *Graphs and Homomorphisms.* Oxford University Press, 2004.

[MRar] Meena Mahajan and B. V. Raghavendra Rao. Small-space analogues of valiant's classes and the limitations of skew formulas. *Computational Complexity*, To Appear.

[Poi08]   Bruno Poizat. À la recherche de la définition de la complexité d'espace pour le calcul des polynômes à la manière de valiant. *Journal of Symbolic Logic*, 73, 2008.

[Sch78]   Thomas J. Schaefer. The complexity of satisfiability problems. In *Proceedings of the tenth annual ACM symposium on Theory of computing*, STOC '78, pages 216–226, New York, NY, USA, 1978. ACM.

[SY10]   Amir Shpilka and Amir Yehudayoff. Arithmetic circuits: A survey of recent results and open questions. *Foundations and Trends in Theoretical Computer Science*, 5(3-4):207–388, 2010.

[Val79a]   L. G. Valiant. Completeness classes in algebra. In *Proceedings of the eleventh annual ACM symposium on Theory of computing*, STOC '79, pages 249–261, New York, NY, USA, 1979. ACM.

[Val79b]   L.G. Valiant. The complexity of computing the permanent. *Theoretical Computer Science*, 8(2):189 – 201, 1979.

[Val01]   Leslie G. Valiant. Quantum computers that can be simulated classically in polynomial time. In *Proceedings of the thirty-third annual ACM symposium on Theory of computing*, STOC '01, pages 114–123, New York, NY, USA, 2001. ACM.

[Wel93]   D.J.A. Welsh. *Complexity: knots, colourings and counting*. London Mathematical Society lecture note series. Cambridge University Press, 1993.