SOME OBSERVATIONS ON HOLOGRAPHIC ALGORITHMS

Leslie G. Valiant

August 1, 2017

Abstract. We define the notion of diversity for families of finite functions, and express the limitations of a simple class of holographic algorithms, called elementary algorithms, in terms of limitations on diversity. We show that this class of elementary algorithms is too weak to solve the Boolean Circuit Value problem, or Boolean Satisfiability, or the Permanent. The lower bound argument is a natural but apparently novel combination of counting and algebraic dependence arguments that is viable in the holographic framework. We go on to describe polynomial time holographic algorithms that go beyond the elementarity restriction in the two respects that they use exponential size fields, and multiple oracle calls in the form of polynomial interpolation. These new algorithms, which use bases of three components, compute the parity of the following quantities for degree three planar undirected graphs: the number of 3-colorings up to permutation of colors, the number of connected vertex covers, and the number of induced forests or feedback vertex sets. In each case the parity can also be computed for any one slice of the problem, in particular for colorings where the first color is used a certain number of times, or where the connected vertex cover, feedback set or induced forest has a certain number of nodes.

Keywords. computational complexity, algebraic complexity, holographic algorithms

Subject classification. 03D15, 15A15, 68Q15, 68Q17, 68R10

1. Introduction

The theory of holographic algorithms is based on a notion of reduction that enables computational problems to be interrelated

with unusual fluidity. The theory offers three basic reduction techniques:

(a) *Holographic transformations* that relate pairs of problems by simply taking a different view or *basis*,

(b) *Holographic gadgets* that use internal cancellations custom designed for the problems at hand, and

(c) *Interpolation techniques* for recovering information from the outputs of computations on a set of specially prepared variants of the problem instance at hand.

The overarching open question in the theory is whether this combination of techniques can bridge the gap between classical polynomial algorithms on the one hand, and the class of #P-complete (or NP- or $\oplus$P-complete) problems as defined by classical reductions, on the other.

In order to further our understanding of this question we introduce here the notion of *diversity* for finite functions, in terms of which some limitations of the simplest kinds of holographic algorithms that we discussed in an earlier paper (Valiant 2006) can be explored more explicitly. These simplest holographic algorithms are those obtained from what we define as *elementary* reductions. We show that such algorithms do impose a limitation on the diversity of the functions that can be realized. It remains unresolved, however, whether holographic algorithms that are not bound by the constraints of elementarity, such as those given in later sections of this paper, that use interpolation, can evade this diversity limitation.

In the later sections we go on to describe some polynomial time holographic algorithms for three natural problems for planar undirected graphs of degree three. These compute the parity of the number of solutions of each of the following three problems: feedback vertex sets (or, equivalently, induced forests), connected vertex covers, and vertex 3-colorings up to permutations of colors. These algorithms use the three element basis **b3** from Valiant (2008). Bases of size greater than two have been studied systematically more recently, for example by Cai *et al.* (2014), Chen (2016), Xia (2016).

For brevity of exposition we shall assume familiarity with the

basic notions and notations of holographic algorithms as described in (Valiant 2004, 2008).

2. Diversity

For a Boolean function $f(x_1, ..., x_m)$ and a subset $S \subseteq X = \{x_1, ..., x_m\}$ of size n, we define the *diversity of* S *in* f to be the logarithm to the base two of the number of different functions of the n variables of S that can be obtained by fixing the $m-n$ remaining variables $X - S$ in the 2^{m-n} different ways. This is the central concept in proofs (Neciporuk 1966) of lower bounds on formula complexity. He showed that if X has a partition into subsets S_i such that the average diversity of the S_i in f is substantial, then a nonlinear lower bound on the formula size of f follows.

We say that a Boolean function $f(x_1, ..., x_m)$ has n-*diversity* D if D is the maximum diversity of S in f, over all subsets $S \subseteq X = \{x_1, ..., x_m\}$ of size n . Since there are 2^{2^n} Boolean functions of n variables, the maximum n-diversity of a function is 2^n.

A Boolean function family $f = \{f_m(x_1, ..., x_m) \mid m = 1, ...\}$ has *diversity* $g(n)$ if for each positive integer n, $g(n)$ is the maximum n-diversity of f_m for any $m \geq n$. Clearly $g(n) \leq 2^n$.

A Boolean function family f has *polynomial diversity* if its diversity $g(n)$ is upper bounded by some polynomial $p(n)$. It has *exponential diversity* if it is lower bounded by $c2^{cn^\kappa}$ for some constants $c, \kappa > 0$. It has *exponential standard diversity* if, for some polynomial $p(n)$, exponential diversity is achieved for all n by f_m with some $m \leq 2^{p(n)}$. It has *polynomial standard diversity* if, for all polynomials $p(n)$, the n-diversity achieved for n by f_m with $m \leq 2^{p(n)}$ is polynomial bounded.

Such definitions can also be made for finite fields F_q for families with f_i: $\{0, ..., q-1\}^i \mapsto \{0, ..., q-1\}$. In that case the maximum n-diversity of a family is $q^n \log_2 q$.

High diversity does not imply high complexity. The Circuit Value problem (Ladner 1975) $\mathrm{CV}_{n,r}(x_1, ..., x_{n+r})$ we shall formulate here as the function that regards its first n inputs as a vector v of n Boolean values, and the remaining bits as a specification of a Boolean circuit C of n inputs with binary gates.

Proposition 1 The circuit value problem has exponential diversity.
Proof An m gate circuit of n inputs can be specified using $r = O((n+m)\log(n+m))$ bits. Since all the 2^{2^n} Boolean functions of n variables can be realized by a circuit with $O(2^n/n)$ gates Lupanov (1958), they can all be encoded in $CV_{n,r}$ if $r = O(2^n)$. We define CV to have such an encoding of circuits. Hence CV has diversity 2^n since $S = (x_1, ..., x_n)$ has diversity 2^n in $CV_{n,r}$ for an appropriate $r = O(2^n)$. Clearly CV then also has exponential standard diversity. □.

Using the following notion of reduction one can deduce that most natural P-, NP- and #P-complete problems have exponential standard diversity. We say that a reduction τ from $\{CV_{n,r}\}$ to a family of functions $\{Q_i\}$ is *segregating* if in polynomial time τ maps the pair v, C to a pair of Boolean sequences (y, z) such that (i) for any fixed n and r, the lengths of y and of z are uniquely determined, (ii) the length of y is polynomially bounded in terms of n, (iii) y depends on v and not C, (iv) z depends on C and not v, and (v) $Q_i(y, z) = C(v)$. (In short, y encodes v, z encodes C, and Q_i evaluates C on v. The length of y is polynomial in n, but the length of z may be exponential in n.)

Proposition 2 *If CV is reducible to Q by a segregating reduction then Q has exponential standard diversity.*
Proof For a fixed size n consider $CV_{n,r}$ with r exponential in n and large enough that all 2^{2^n} Boolean functions of n variables can be expressed. Now consider *one* of the 2^{2^n} choices of C. Since the reduction, say τ, is segregating, for all v it will map (v, C) to (y, z) for the some fixed value of z. For C and z so fixed, as v varies so will y, and $Q_i(y, z) = C(v)$. Hence, Q_i will compute on the encoding y of v the same Boolean function as C does on v. Hence, fixing z in different ways will make Q_i compute 2^{2^n} different functions of y. If S is the set of variables that represents y then the diversity of S in Q_i will be 2^n. Since τ is segregating, by condition (ii) $|S|$ is polynomially bounded in terms of n. It follows that the diversity of S in Q_i will be at least $c2^{c|S|^\kappa}$ for appropriate positive constants

c and κ. □.

Now for many NP-complete problems, by tracing through the known reductions, one can derive segregating reductions from CV to them. For example, consider the family Q corresponding to Cook's 3SAT problem. Here $Q_i(x)$ is a 3CNF formula with i clauses and variables from $x_1, ..., x_i$. From a circuit C with inputs $x_1, ..., x_n$, and any vector v of values of $x_1, ..., x_n$, one can construct by now standard methods a polynomial size 3CNF formula that is satisfiable if and only if that circuit C on that input v evaluates to one: The formula will have the first n clauses encode the input with the j^{th} clause being (x_j) or (x'_j) according to whether the j^{th} among the n bits of v is 1 or 0. It will have the remaining clauses encode the gates. This is a segregating polynomial time reduction from CV to 3CNF. Related to the 3CNF satisfiability problem is $\oplus$3CNF, the problem of determining the parity of the number of solutions of a 3CNF formula, and their planar analogs Pl-3CNF and $\oplus$Pl-3CNF. From the above construction we can deduce the following.

Proposition 3 *The problems 3CNF, $\oplus$3CNF, Pl-3CNF and $\oplus$Pl-3CNF all have exponential standard diversity.*
Proof The previous paragraph describes a segregating reduction from CV to 3CNF. This establishes the result for 3CNF by virtue of Proposition 2. Since the construction can be made to preserve the number of solutions, the 3CNF formula will have 0 or 1 solutions according to whether the value output by the circuit C is 0 or 1. The result for $\oplus$3CNF therefore also follows. For the planar case one uses additional sets of clauses that act as crossovers and make the formula planar, as described by Lichtenstein (1982). These can also be made to preserve the number of solutions (Hunt *et al.* 1998). These additional clauses can be viewed as part of the circuit encoding, and then yield a segregating reduction to the planar versions of Pl-3CNF and $\oplus$Pl-3CNF as needed. □

With this starting point one can ask for each of the known NP-complete problems, such as those of Karp (1972), whether CV

is reducible to some natural encodings of them by a segregating reduction. It appears that this is the case for the vast majority, and for those it then follows that some natural encoding of them has exponential standard diversity.

What is the status of the numerous counting problems that are known to be complete in an appropriate counting class, but for which existence is polynomial time computable and not known to be complete for P? Do these counting problems have CV embedded in them equally explicitly? The following shows that in some such cases the embedding is in fact explicit.

Proposition 4 *The Permanent modulo k for any prime $k \neq 2$ has exponential standard diversity.*
Proof From the proof in Valiant (1979a) one can obtain a segregating reduction from CV via 3CNF to the permanent modulo k for any prime k other than two. □

On the other hand, proofs of #P-completeness often go through interpolation (Cai *et al.* 2008; Jerrum 1987; Vadhan 2001; Valiant 1979b; Xia *et al.* 2007). It is an open problem whether in those cases exponential diversity is necessarily implied. For example, does counting matchings modulo 3, in some natural encoding of planar graphs, have exponential diversity? There are cases in which the known reductions to the counting problem take particularly circuitous routes through interpolations, raising the possibility that the CV problem is truly disguised, but nevertheless exponential diversity can be deduced from known reductions for the corresponding parity problem. One example of this is planar vertex cover for which known #P completeness proofs (Vadhan 2001; Xia *et al.* 2007) are indirect. However, for the subclass of planar regular 3/2-graphs (bipartite graphs with degree 2 on one side and 3 on the other) a segregating reduction from 3CNF to this vertex cover problem that preserves the parity of the number of solutions can be derived from the the $\oplus$P-completeness proof of this problem given in Valiant (2006).

Proposition 5 *The parity of the number of vertex covers for pla-*

nar regular 3/2 graphs has exponential standard diversity.

Clearly a unary Boolean function family, one that is zero whenever any $x_i = 1$, will have polynomial diversity. Since there exist unary function families of arbitrarily high Turing machine time complexity it follows that polynomial diversity does not imply polynomial time Turing machine complexity. Pavel Pudlák has made the following elegant observation that shows that low n-diversity can also be possessed by functions that have high complexity in many other senses, such as having exponential circuit complexity or being NP-complete. For any function $f(x)$ on n inputs consider an error correcting code g: $\{0,1\}^n \rightarrow \{0,1\}^r$ that corrects more than n errors and can be computed and inverted efficiently. Let $h : \{0,1\}^r \rightarrow \{0,1\}$ be such that $h(y) = 1$ iff $y = g(x)$ for some x and $f(x) = 1$. Then h has n-diversity at most $n + 1$ since for any domain d of n of its input bits, fixing the remaining bits will permit it to have value 1 for at most 1 of the 2^n values of d, and hence there are at most $2^n + 1$ such different functions possible of the n variables of d.

In other words, functions of arbitrary difficulty can be made to have low n-diversity for n appropriately chosen. However, the above discussion does not preclude the function h having exponential n' diversity for some $n' \geq n$. Indeed the following appears to be unresolved.

Open Problem Does any function family f that is NP-complete, $\oplus$P-complete, #P-complete, or PSPACE-complete, have polynomial diversity?

3. Elementary Reductions to Matchgrids

We shall now define the notion of an *elementary reduction to matchgrids.* The definition is a generalization of the one given in Valiant (2006) that was specific to reductions from 3CNF. Here we consider a *Boolean function family f with respect to input domain d.* for each m there is a specified function f_m on Boolean variables $x_1, ..., x_m$, and a specified subset d_m of n_m of these m variables.

Such a family, like 3CNF, is assumed to have a natural representation, and a natural representation size that is robust to polynomial factors.

We say that τ is a $k(m)$-*oracle reduction from family f to matchgrids* if for each family member f_m it generates $k(m)$ matchgrids in polynomial time and from their Holants it computes the solution to the original problem also in polynomial time. Note that while many holographic algorithms in the literature are 1-oracle, several multi-oracle matchgrid reductions that use interpolation have been described also (Cai & Choudhary 2007; Valiant 2004, 2008).

Suppose that τ is a polynomial time 1-oracle reduction from f to matchgrids over field F. For a function f_m, a domain of its variables d_m of size n_m, and an assignment z to the set c_m of variables that is the complement of d_m, let $\mathrm{M}(f_m, d_m, z)$ be the set of 2^n adjacency matrices of the set of matchgrid images under τ of the 2^n restrictions of f_m when the $n = n_m$ variables specified by d_m are fixed in all possible ways.

Then τ is a *local boundary reduction for family f and domain d* if for each m, z the matchgrids produced have adjacency matrices $\mathrm{M}(f_m, d_m, z)$ and planar embeddings such that
(a) the 2^n embeddings have an identical set of nodes $U_{m,z}$.
(b) in these 2^n embeddings all the edges and their weights are identical, except possibly those that have both endpoints within a subset $Z_{m,z} \subseteq U_{m,z}$, which is of size upper bounded by a polynomial $L(n)$ independent of m.
(c) the $Z_{m,z}$ nodes have degrees bounded by a constant independent of n, m or z.
(d) the nodes $Z_{m,z}$ all lie in the infinite outer face of the embedding of the graph induced by $T_{m,z} = U_{m,z} - Z_{m,z}$, and
(e) the edges incident to pairs of vertices in $Z_{m,z}$ can be partitioned into n sets such that each such set S_i corresponds to a variable x_{k_i} in d_m, and the weights of S_i are functions of the value of x_{k_i} but are independent of the values of the other x_j in d_m (i. e. those with $j \neq k_i$.)

We shall say that *a reduction τ from a family f with respect to input domain d of size n to matchgrids over F* is *elementary* if it has the four properties of (i) being 1-oracle, (ii) being local boundary,

(iii) having the number of field elements $|F|$ polynomial bounded in m, and (iv) having Holant$(\tau(f(x)))$ determine the value of $f(x)$. We then also say that τ is an *elementary reduction for* (f, d).

We note that there are no constraints on what the transformation does on different z. The intent is that when the Circuit Value problem is embedded, then the different circuits C can be embedded in arbitrarily different ways, but for any one circuit there are constraints on the way the matchgrids can vary as the inputs v to C vary. Also note that the field size is allowed to grow polynomially with m, so that it can be exponential in n if m is exponential in n.

4. Elementary Reductions Compute Functions of Polynomial Diversity

Here we provide a lower bound argument that proves limitations on the functions that can be obtained as elementary reductions to matchgrids. The argument is a simple combination of counting and algebra. The counting part dictates the restriction to fields of limited size. The algebraic part captures the notion that the different components of a signature are Pfaffians of submatrices of a common matrix and therefore obey some algebraic relationships that limit the possible signatures.

We start with a universality statement that all realizable signatures having a fixed number of components are realizable by a single matchgate with parameters set to certain values. A prototype of such a result for the related context of *matchcircuits* and *characters* was proved for 2-input 2-output gates in Valiant (2002). This was subsequently generalized by Cai *et al.* (2009) and Cai & Gorenstein (2014) to arbitrary size gates and shown to be the essentially equivalent problem for the planar *matchgrids* and *signatures* considered in Valiant (2004) and here.

Theorem 1 *There is a weighted planar graph G having r external nodes and $O(r^4)$ edges of which all but $O(r^2)$ have fixed weight ± 1, such that for any field F, any standard signature that is realized by some matchgrid with r external nodes can be realized by G by*

setting the $O(r^2)$ variable-weight edges to appropriate constants.
Proof This follows most directly from Cai & Gorenstein (2014). Their Theorem 7 provides such a construction for matchcircuits. It consists of a planar realization of a complete graph with crossover gadgets at the intersections. The general result then follows from the equivalence of matchgrids and matchcircuits also shown there. □

Theorem 2 *For any family f and domain d if there is an elementary reduction to matchgrids for (f, d), then d has polynomial standard diversity in family f.*
Proof Suppose that τ is a polynomial time 1-oracle reduction from f to matchgrids over field F. For a function f_m and a domain of its variables d_m of size $n = n_m$, for each z let $\mathrm{M}(f_m, d_m, z)$ be the set of adjacency matrices as defined above and consider the planar embeddings that respect conditions (a)-(e) of the definition of elementarity.

By (a), (b) these embeddings are identical with respect to all the edges that are incident to a $T_{m,z}$ node at least at one end . We regard the embedding of the nodes $T_{m,z}$ as a matchgrid $H_{m,z}$. By (d) the remaining nodes $Z_{m,z}$ are all mapped into the outer face of $H_{m,z}$. Since, by (b) and (c), $|Z_{m,z}|$ is upper bounded by a polynomial $L(n)$, and the degrees of the $Z_{m,z}$ nodes by a constant, $H_{m,z}$ has $\mathrm{O}(L(n))$ external connections, and can be regarded as a matchgrid with $\mathrm{O}(L(n))$ external nodes. Now, by Theorem 1, $H_{m,z}$ can be replaced by a matchgrid with $\mathrm{O}((L(n))^2)$ variable weight edges. From this we deduce that as z varies, the total number of inequivalent matchgrids $H_{m,z}$ is at most $a_1 = |F|^{O((L(n))^2)}$. (In other words it is single exponential in n however large m may be.)

It remains to complete the estimation of the number of different functions of the original n domain d_m variables that the matchgrids can realize as z varies, by also taking into account the remainder of the matchgrid specification, namely the nodes $Z_{m,z}$ and the edges incident to them. We can fix the names of the nodes of $Z_{m,z}$ and the external nodes of $H_{m,z}$, which altogether number $\mathrm{O}(L(n))$. Then the number of potential edges that have at least one endpoint in $|Z_{m,z}|$ is at most $A = \mathrm{O}((L(n))^2)$. By assumption (e), each choice

of z partitions the edges among pairs of $Z_{m,z}$ nodes into n sets, and each such set will have a weight assignment that represents the corresponding domain variable having value zero, and a weight assignment corresponding to value one. The number of partitions is upper bounded by $a_2 = n^A$ clearly. Also, for each such partition, $a_3 = |F|^{2A}$ upper bounds the number of distinct weightings of the edges between $Z_{m,z}$ nodes that among them represent all combinations of 0's and 1's for the n variables of d_m. Also, the number of possible weight assignments to edges incident to both $Z_{m,z}$ and $T_{m,z}$ nodes is upper bounded by $a_4 = |F|^{O(L(n))}$. It follows that the total number of functions of the domain variables that the matchgrids can realize is upper bounded by $a_1a_2a_3a_4$, which itself is upper bounded by $(nL(n)|F|)^{O((L(n))^2)}$. Now, by condition (iii) of elementarity, $|F|$ is polynomial bounded in m. For standard diversity m is single exponential in a polynomial $p(n)$. It follows that the number of distinct functions is at most $2^{O(q(n))}$ for some polynomial q. In other words the standard diversity is at most polynomial in n. □

From this result one can deduce for problems known to have high diversity that they do not have elementary reductions to matchgrids. The following is an instance that parallels a result in Valiant (2006):

Corollary 1 There is no elementary reduction from (f, d) to matchgrids where f is any one of Pl-3CNF, 3CNF, $\oplus$Pl-3CNF or $\oplus$3CNF, and d specifies a subset of $\mathrm{O}(\log m)$ of the clauses for formulae with m clauses.
Proof This follows from Proposition 3 and Theorem 2. □

Note that this corollary implies that one should not expect that the planar Circuit Value Problem with inputs on the periphery can be mapped into a matchgrid by an elementary reduction.

It is an interesting question whether implications of a converse nature also hold. For problems such as $\#_7$Pl-Rtw-Mon-3CNF (Valiant 2006) for which 1-oracle holographic algorithms exist, even for fields whose size does not increase at all with the input size, one would like to determine whether they have polynomial diver-

sity. For planar representations of Boolean functions one can define a notion of *planar diversity* where the domains d_m have to be on the outer face of the embedding. Then the particulars of the algorithm just described for $\#_7$Pl-Rtw-Mon-3CNF do imply polynomial planar diversity for that problem. However, such arguments do not appear to apply to domains that are not on the periphery, or to general diversity.

Multiple oracle calls appear to be very useful in reductions among counting problems. There are multitudes of #P-complete problems that have been proved complete via reductions that involve multiple oracle calls and polynomial interpolation on the results (Cai *et al.* 2008; Jerrum 1987; Vadhan 2001; Valiant 1979b; Xia *et al.* 2007). For any one of these problems one can ask whether they have polynomial diversity.

In the opposite direction, one can ask whether algorithms that make multiple oracle calls, each via an elementary reduction, can compute functions of exponential diversity. To formulate specific questions of this kind one would need to define specific classes of such multiple oracle call algorithms. One relevant such class is offered by the algorithms described in Sections 6-8 of this paper. These all have the following form: Given an instance G of the problem, one generates a single matchgrid with weights that are polynomials in x with coefficients from a field F. The solution sought is the jth least significant bit in the coefficient of x^i of the Holant, where i, j are predetermined integers, and all the coefficients are guaranteed to be integral. There is the further constraint that if this coefficient is nonzero then it has at least $j-1$ factors of 2. Note that the solutions here are obtained by the multi-oracle reduction that evaluates the matchgrid at enough different values of x, and then interpolates for the appropriate coefficient. It is an interesting question to determine whether or not these classes of reductions can evade the polynomial constraint on diversity of elementary reductions.

5. The basis b3

The basis **b3** (Valiant 2008) has three components $\mathbf{z} = (1, 0)$, $\mathbf{n} = (1, -1)$, $\mathbf{p} = (1, 1)$. It has the useful property that for all $x \in F$,

$x\mathbf{z}^3 + \mathbf{n}^3 + \mathbf{p}^3$ is an even ternary signature and therefore, by Proposition 6.2 in Valiant (2008), is realizable by a planar matchgate. To verify this it is sufficient to expand $x\mathbf{z}^3 + \mathbf{n}^3 + \mathbf{p}^3$ as :

$(x,0,0,0,0,0,0,0)+(1,-1,-1,1,-1,1,1,-1)+(1,1,1,1,1,1,1,1)$
$= (x+2,0,0,2,0,2,2,0) = [x+2,0,2,0]$.

We shall call this signature, and the gate realizing it, $\mathbf{g}_3(x)$. The analogous two-output signature $\mathbf{g}_2(x)$ is also even, and therefore realizable by virtue of Proposition 6.1 in Valiant (2008), since

$x\mathbf{z}^2 + \mathbf{n}^2 + \mathbf{p}^2 = (x,0,0,0) + (1,-1,-1,1) + (1,1,1,1) = [x+2,0,2]$,

as is also the one output signature $\mathbf{g}_1(x) = x\mathbf{z} + \mathbf{n} + \mathbf{p} = [x+2,0]$.

For each of the three parity problems that we define in the sections that follow, we shall consider planar graphs of n vertices all of maximum degree three. Our constructions do not require that the graph be *cubic* in the sense that every node has degree exactly three.

For each problem we shall construct for any such graph G a family of matchgrids $\Omega(G,x)$ indexed by x, using a fixed binary recognizer $\mathbf{r}$ for the edges, and the above mentioned generators $\mathbf{g}_1(x)$, $\mathbf{g}_2(x)$ and $\mathbf{g}_3(x)$, for the nodes of degrees one, two and three, respectively. Then for each problem, Holant$(\Omega(G,x))$ can be viewed as a polynomial in x of degree at most n. If we evaluate Holant$(\Omega(G,x))$ for one G and $n+1$ distinct values of x, and interpolate for the coefficients, then the coefficient of x^i will be the sum of the contributions to the Holant of the states in which exactly i of the generators are generating $\mathbf{z}$'s, and the remainder $\mathbf{n}$'s or $\mathbf{p}$'s.

Alternatively, we shall sometimes substitute $\mathbf{g}_1(s)$, $\mathbf{g}_2(t)$ and $\mathbf{g}_3(x)$, with different indeterminates s,t,x. Then after evaluating at $\mathrm{O}(n^3)$ distinct points, we can interpolate to obtain the coefficient of $s^i t^j x^l$, which gives the contribution to the Holant of states where among the $\mathbf{z}$ generators, exactly i have degree one, j degree two, and l degree three.

We now describe the binary recognizers that we use for the edges. Each of these recognizers is a simple chain, of one or two edges, with the end nodes serving as the two external nodes. In our notation below * denotes a node, and *(w)* denotes an edge

of weight w between two nodes. The following can be verified by inspection.

Proposition 6 *The values of the following three recognizers are as follows when* (a, b) *is input from the left, and* (c, d) *from the right:*
$\mathbf{r}_1$*:* $*(1)*$ *has value* $ac + bd$*,*
$\mathbf{r}_2$*:* $*(1)*(1)*$ *has value* $ad + bc$*, and*
$\mathbf{r}_3$*:* $*(1)*(\text{-}1)*$ *has value* $ad - bc$*.*

6. Holographic Algorithm for the Parity of the Number of Induced Forests or Feedback Vertex Sets

The Minimum Feedback Vertex Set problem for undirected graphs is defined as follows: Given an undirected graph G and an integer k the question is to determine whether there is a set of k vertices whose removal leaves a forest (i.e. a graph with no cycles.) There is a substantial literature on this existence problem. The directed version of this problem was proved NP-complete by Karp (1972). The undirected version we study here was proved NP-complete by Garey & Johnson (1979). Subsequently it was shown to be NP-complete even for planar graphs of degree four by Speckenmeyer (1983). For cubic (i.e. regular degree three) graphs a polynomial algorithm was given by Li & Liu (1999). (This last result is also implied by the polynomial time algorithm of Ueno *et al.* (1988) for the Minimum Connected Vertex Cover problem (defined in the next section) for cubic graphs, in conjunction with the result of Speckenmeyer (1983, 1988) that for any cubic graph on n vertices MCVC-MFVS $= n/2 - 1$, where MCVC and MFVS denote the sizes of the minimum connected vertex cover and the minimum feedback vertex set.)

Here we are interested not in the existence problem but in the parity of the number of solutions, not only for forests of the largest size but for forests of every size, and not only for regular graphs of degree three, but for all graphs of maximum degree three. However, we restrict ourselves here to planar graphs. Thus the problem we address, $\oplus PlmFVS$, is the following: Given a degree m planar

undirected graph G and an integer k, determine the parity of the number of sets of k nodes that induce a forest in G.

Theorem 3 *There is a deterministic polynomial time algorithm for $\oplus$Pl3FVS.*

Proof We place $\mathbf{g_3(x)}$, $\mathbf{g_2(x)}$ and $\mathbf{g_1(x)}$ generators at vertices of degree three, two and one respectively. We place a recognizer $\mathbf{r}_1$ on each edge. Then, by Proposition 6, the value of each recognizer as a function of the nine possible combinations of what the adjacent nodes generate are as follows: $\mathbf{zz} \to 1$; $\mathbf{zp} \to 1$; $\mathbf{zn} \to 1$; $\mathbf{pz} \to 1$; $\mathbf{nz} \to 1$; $\mathbf{pp} \to 2$; $\mathbf{nn} \to 2$; $\mathbf{pn} \to 0$; $\mathbf{np} \to 0$.

We regard each state σ (i.e. each combination of states of all the generators) of the matchgrid as a two-coloring, where one color, Z, corresponds to the nodes generating **z**'s, and the other, Y, those generating **n**'s and **p**'s. For each such state we define $\#YY(\sigma)$ to be the number of edges joining a pair of nodes both colored Y, and $\#Ycomponents(\sigma)$ to be the number of connected components induced in G by the removal of the Z nodes and the edges adjacent to them. Then the Holant will be the sum over all such Z/Y 2-colorings of G of the value $U = 2^{\#Ycomponents(\sigma)+\#YY(\sigma)}$, since each connected component has one of two states (all **n** or all **p**), and each edge in such a component contributes a further factor of two. If G has n nodes and the number of Z nodes is fixed as $n-k$, then the minimum number of divisors of 2 in U is $2^{n-(n-k)} = 2^k$, and is achieved if and only if the YY edges induce a forest in G. (Note that in any graph with k nodes the sum of the number of edges and the number of connected components is at least k, the minimum being achieved only if the graph is a forest.) Hence, if one divides the coefficient of x^{n-k} in Holant$(\Omega(G, x))$ by 2^k, then the parity of that number is the desired solution to $\oplus$Pl3FVS(G, k). □

7. Holographic Algorithm for the Parity of the Number of Connected Vertex Covers

The Minimum Connected Vertex Cover problem is the following. Given an undirected graph G determine the size of the smallest set of nodes that (i) is a vertex cover, and (ii) induces a connected

subgraph of G.

The existence problem was shown NP-complete for degree four planar graphs by Garey & Johnson (1977). Fernau & Manlove (2009) showed that this result holds even in the bipartite case. For cubic graphs it was shown to be polynomial time computable by Ueno *et al.* (1988).

Here we are interested in the following parity problem $\oplus PlmCVC$. Given an undirected planar graph G of maximum degree m and an integer k, the problem is to compute the parity of the number of connected vertex covers of G of k vertices.

Theorem 4 *There is a deterministic polynomial time algorithm for* $\oplus Pl3CVC$.

Proof We place $\mathbf{g_3(x)}$, $\mathbf{g_2(t)}$, and $\mathbf{g_1(s)}$ generators at vertices of degree three, two and one respectively. We place a recognizer $\mathbf{r}_2$ on each edge. Then, by Proposition 6, the value of each recognizer as a function of the nine possible combinations of what the adjacent nodes generate are as follows: **zz** $\rightarrow$ 0; **zp** $\rightarrow$ 1; **zn** $\rightarrow$ -1; **pz** $\rightarrow$ 1; **nz** $\rightarrow$ -1; **pp** $\rightarrow$ 2; **nn** $\rightarrow$ -2; **pn** $\rightarrow$ 0; **np** $\rightarrow$ 0.

As before, we regard each state σ of the matchgrid as a two coloring, where one color, Z, corresponds to the nodes generating **z**'s, and the other, Y, those generating **n**'s and **p**'s. For each such state we define #YY(σ) to be the number of edges joining a pair of nodes both colored Y, and #Ycomponents(σ) to be the number of connected components induced in G by these YY edges. Now the Holant will be the sum, over some such Z/Y 2-colorings of G in which the nodes colored Z form an independent set, of $U = \pm 2^{\#Ycomponents(\sigma)+\#YY(\sigma)}$. This will follow by a similar argument to that used in Theorem 3, except now the Z nodes form an independent set since the value of $\mathbf{r}_2$ for **zz** input is zero, and we need to analyze potential cancelations.

We first consider the case that the graph has an even number of edges. To derive this value of U we first note that if the graph has n nodes and is cubic, then for a state in which the Z nodes form an independent set of size $n-k$, it will be the case that $\#YY(\sigma) = 3n/2 - 3(n-k) = 3(k-n/2)$. For each Y/Z coloring and for any connected component induced by the Y colored edges

in G, there will be two valid states, corresponding to the Y-colored nodes having all **p** or all **n** states. When one changes all the Y nodes from **p** to **n** then the values of all the recognizers in G will change sign. Hence if the nodes in this component have an even number of edges incident to them in G then these contributions to the Holant will have the same sign, and otherwise will cancel. Hence the minimum (nonzero) number of divisors of 2 in U is $3(k - n/2) + 1$, and is achieved if and only if the YY edges induce one connected component in G and G has an even number of edges. Hence, if one divides the coefficient of x^{n-k} in Holant$(\Omega(G, x))$ by $2^{3(k-n/2)+1}$, then the parity of that number is the desired solution to $\oplus$Pl3CVC(G).

If the graph is not regular, then by interpolation we can find the coefficient of $s^i t^j x^l$ in Holant$(\Omega(G, s, t, x))$ for all i, j, l. For any specific combination of i, j, l the value of $\#\mathrm{YY}(\sigma)$ is $|E|-i-2j-3l$, where $|E|$ is the total number of edges in G. Hence we can compute the parity of the number of solutions for any combination i, j, l, and hence for all the combinations with $i+j+l = n-k$. We shall derive the parity of the number of solutions corresponding to such Z sets by dividing the appropriate coefficient by $2^{|E|-i-2j-3l+1}$ rather than by $2^{3(k-n/2)+1}$ as used in the regular case.

So far we have assumed that the number of edges in G is even. To treat the alternative case we choose an arbitrary edge and replace $\mathbf{r}_2$ by $\mathbf{r}_1$ on it. This ensures that when switching between all **p** and all **n** states the sign will not change on this one edge, and hence not for the product of all of these odd number of edges. It only remains to ensure that the Y nodes still form a vertex cover, and for this it is necessary to preclude that the endpoints of the chosen edge be both in state Z. This can be done by multiplying the x term in these two generators by a new indeterminate w, and, by interpolation, computing and adding the coefficients of w^0 and w^1 (while ignoring that of w^2). □

8. Holographic Algorithm for the Parity of the Number of Vertex Colorings

A 3-Vertex Coloring of a graph G is an assignment of a color from a palette of 3 colors to each vertex so that no pair of adjacent vertices has the same color. Clearly the set of all such proper colorings can be partitioned into equivalence classes of 3! colorings, so that the members of each class differ only by a permutation of the colors. Here we are interested in the following two closely related problems. The problem $\oplus PlmCol$: for an undirected planar graph G of maximum degree m determine the parity of the number of equivalence classes of 3-colorings of G. The problem $\oplus PlmFCol$: for an undirected planar graph G of maximum degree m and an integer k determine the parity of the number of 3-colorings that are invariant under permutation of the second and third color, when exactly k nodes are given the first color. We note that the corresponding counting problems for 3-colorability of degree three graphs are #P-complete Bubley *et al.* (1999).

Theorem 5 (Barbanchon 2004) *For some constant m, $\oplus PlmCol$ is $\oplus$P-complete.*

Theorem 6 *There is a deterministic polynomial time algorithm for $\oplus Pl3FCol$ and for $\oplus Pl3Col$.*

Proof We place $\mathbf{g_3(x)}$, $\mathbf{g_2(t)}$, and $\mathbf{g_1(s)}$ generators at vertices of degree three, two and one respectively, and $\mathbf{r}_3$ recognizers on each edge. The $\mathbf{r}_3$ recognizers for $ad-bc$ are not symmetric, and can be placed in arbitrary orientation without influencing our result. By Proposition 6 the value of each recognizer as a function of the nine possible combinations of what the adjacent nodes generate are as follows: **zz** $\rightarrow$ 0; **zp** $\rightarrow$ 1; **zn** $\rightarrow$ -1; **pz** $\rightarrow$ -1; **nz** $\rightarrow$ 1; **pp** $\rightarrow$ 0; **nn** $\rightarrow$ 0; **pn** $\rightarrow$ -2; **np** $\rightarrow$ 2.

Again we regard each state σ of the matchgrid as a two coloring, where one color, Z, corresponds to the nodes generating **z**'s, and the other, Y, those generating **n**'s and **p**'s. For each such state we define #YY(σ) to be the number of edges joining a pair of nodes both colored Y, and #Ycomponents(σ) to be the number of connected components induced in G by these YY edges. Then the Holant

will be the sum over some such Z/Y 2-colorings of G in which the nodes colored Z form an independent set and those colored Y form a bipartite graph, of the values $U = \pm 2^{\#Ycomponents(\sigma)+\#YY(\sigma)}$.

To see this we first assume that the graph is cubic and has an even number of edges. If the graph has n nodes, then for a state in which the Z nodes form an independent set of size k, then $\#\mathrm{YY}(\sigma) = 3n/2 - 3k$. We note that for each Y/Z coloring the YY edges will form a set of connected bipartite components in G. In each component there will be two valid states, corresponding to which of the two parts is in **p** or **n** state. When one swaps **p** and **n** all the values of all the recognizers will change sign. Hence if there are an even number of edges incident to the nodes in one such component, then the contributions to the Holant will have the same sign for the two states. Hence the minimum number of divisors of 2 in U is $3n/2-3k+1$, and is achieved if and only if the YY edges induce one connected bipartite component in G. Hence, if one divides the coefficient of x^k in $\mathrm{Holant}(\Omega(G, x))$ by $2^{3n/2-3k+1}$, then the parity of that number is the parity of the number of solutions to $\oplus\mathrm{Pl3FCol}(G, k)$. The sum of these, modulo 2, for $k = 1, ..., n$, is the solution to $\oplus\mathrm{Pl3Col}(G)$, since each solution to $\mathrm{Pl3Col}(G)$ corresponds to three solutions of $\mathrm{Pl3FCol}(G, k)$ instances.

Graphs that are not regular can be treated exactly as in Theorem 4. Graphs with an odd number of edges can be treated by picking one boundary edge and simulating its effect as if all entries in the signature of that recognizer were nonnegative but unchanged in magnitude. This can be done using the fact that any signature for gates with one external node can be realized by boundary gates. (See Valiant (2008). This also holds for two external nodes.) Then if the chosen boundary edge has both endpoints of degree three, for example, we would compute the Holant for the 9 matchgrids obtained by replacing the left generator $\mathbf{g_3(x)}$ by generators for each of **z**, **p**, and **n**, in turn, and the right generator $\mathbf{g_3(x)}$ by the same choice, and combining the 9 values of the resulting Holants to reflect the intended signature component of the $\mathbf{r}_3$ recognizer that is being replaced. □

Following a suggestion of a referee, we note that in this section

we could have eliminated 2's from the signature by using the basis $\mathbf{z} = (s^{-1}, 0)$, $\mathbf{n} = (s, -s)$, $\mathbf{p} = (s, s)$, where $s = 2^{-\frac{1}{2}}$.

For completeness we mention some applications to the edge coloring problem. Here we are counting valid edge colorings, again modulo permutations of colors. We defined planar regular 3/2-graphs earlier. Here we define planar 3/2-graphs to be bipartite graphs where one side has degree at most three and the other side degree at most 2.

Theorem 7 *The parity of the number of edge 3-colorings of planar 3/2 graphs can be computed in polynomial time.*
Proof The line graph of such a graph is a planar degree three graph, and hence the result follows immediately from Theorem 6. □

The following states equivalences between two pairs of problems whose complexities are currently unresolved. Note that in naming problems we use cubic to denote regular graphs of degree three, and the number 3 to refer to graphs of maximum degree three. HC is the Hamiltonian circuit problem. Note that as far as their counting analogues, #PlCubicHC was proved #P-complete in Liśkiewicz *et al.* (2003), and #PlCubicEdgeColor was so proved in Cai *et al.* (2014).

Theorem 8 $\oplus$*CubicHC is polynomial time reducible to* $\oplus$*CubicEdgeColor. Similarly* $\oplus$*PlCubicHC is polynomial time reducible to* $\oplus$*PlCubicEdgeColor.*
Proof Pick any vertex v and fix the three colors, say Y, R and B, arbitrarily for the three incident edges. In any coloring of G we call a maximal set of edges that are connected and each colored Y or R a *YR-component*. Note that any such component must be a simple cycle. Any coloring that contains $i > 1$ YR-components belongs to an equivalence class of 2^{i-1} colorings that have the same pattern of B colors, and differ only in the coloring of the YR-components, except for the one that includes the chosen vertex v where the colors are fixed. Hence, the only B patterns that contribute an odd number of colors are those that have exactly one YR-component,

or equivalently, one cycle including all the vertices. □

For the problem of edge coloring planar graphs of degree at most three there are some formulations that can be shown to be ⊕P-complete by reduction to ⊕Pl3HC which was shown to be ⊕P-complete in Valiant (2005). Let Pl3EdgeColorYR be the problem of edge coloring a planar graph where every vertex has degree at most three, and the edge coloring is restricted so that every edge adjacent to a vertex of degree less than three must have color Y or R.

Theorem 9 *⊕Pl3EdgeColorYR is ⊕P-complete.*
Proof The argument of the proof of Theorem 8 still applies: any YR-component must be a simple cycle, and the parity of the number of such colorings is ⊕Pl3HC the parity of the number of Hamiltonian cycles, which is ⊕P-complete for planar degree three graphs, as observed above. □

However, ⊕3EdgeColor and ⊕Pl3EdgeColor appear to be open.

9. Acknowledgements

A preliminary version of this paper was published as *Proc. 9th Latin American Theoretical Informatics Symposium, LATIN 2010*: Oaxaca, Mexico, April 19-23, LNCS, Vol 6034 Springer-Verlag (2010), 577-590. This research was supported in part by National Science Foundation grants CCF-09-64401 and CCF-15-09178.

References

R. Barbanchon (2004). On unique graph 3-colorability and parsimonious reductions in the plane. *Theoretical Computer Science* **319**(1-3), 455–482.

R. Bubley, M. Dyer, C. Greenhill & M. Jerrum (1999). On approximately counting colourings of small degree graphs. *SIAM J. Comput* **29**, 387–400.

J-Y. Cai & V. Choudhary (2007). Some Results on Matchgates and Holographic Algorithms. *International Journal of Software and Informatics* **1**, 1.

J.-Y Cai, V. Choudhary & P. Lu (2009). On the Theory of Matchgate Computations. *Theory of Computing Systems* **45**(1), 108–132.

J.-Y. Cai & A. Gorenstein (2014). Matchgates Revisited. *Theory of Computing* **10**, 167–197.

J.-Y. Cai, H. Guo & T. Williams (2014). The Complexity of Counting Edge Colorings and a Dichotomy for Some Higher Domain Holant Problems. *FOCS* 601–610.

J.-Y. Cai, P. Lu & M. Xia (2008). Holographic Algorithms by Fibonacci Gates and Holographic Reductions for Hardness. *FOCS* 644–653.

S. Chen (2016). Basis collapse for holographic algorithms over all domain sizes. *STOC* 776–789.

H. Fernau & D. Manlove (2009). Vertex and edge covers with clustering properties: Complexity and algorithms. *Journal of Discrete Algorithms* **7**(2), 149–167.

M. R. Garey & D. S. Johnson (1977). The rectilinear Steiner tree problem is NP complete. *SIAM Journal of Applied Mathematics* **32**, 826–834.

M. R. Garey & D. S. Johnson (1979). *Computers and Intractability: A Guide to the Theory of NP-Completeness.* W. H. Freeman.

H. B. Hunt, M. V. Marathe, V. Radhakrishnan & R. E. Stearns (1998). The Complexity of Planar Counting Problems. *SIAM J. Comput* **27**(4), 1142–1167.

M. R. Jerrum (1987). Two-dimensional monomer-dimer systems are computationally intractable. *J. Statist. Phys* **48**(1-2), 121–134.

R. M. Karp (1972). Reducibility among combinatorial problems. In *Complexity of Computer Computations*, R. E. Miller & J. W. Thatcher, editors, 85–104. Plenum Press.

R. E. Ladner (1975). The Circuit Value Problem is Log Space Complete for P. *SIGACT NEWS* **7**(1), 18–20.

D. M. Li & Y. P. Liu (1999). A polynomial algorithm for finding the minimum feedback vertex set of a 3-regular simple graph. *Acta Math. Sci* **19**(4), 375–381.

D. Lichtenstein (1982). Planar formulae and their uses. *SIAM J. Comput* **11**, 329–343.

M. Liśkiewicz, M. Ogihara & S. Toda (2003). The complexity of counting self-avoiding walks in subgraphs of two-dimensional grids and hypercubes. *Theoretical Computer Science* **304**, 1.

O. B. Lupanov (1958). A method of circuit synthesis. *Izv. VUZ Radiofiz* **1**, 120–140.

E. I. Neciporuk (1966). A Boolean Function. *Sov. Math. Dokl* **7**, 999–1000.

E. Speckenmeyer (1983). *Untersuchungen zum Feedback Vertex Set Problem in ungerichteten Graphen.* Ph.D. thesis, Universität Paderborn.

E. Speckenmeyer (1988). On feedback vertex sets and nonseparating independent sets in cubic graphs. *Journal of Graph Theory* **12**(3), 405–412.

S. Ueno, Y. Kajitani & S. Gotoh (1988). On the nonseparating independent set problem and feedback set problem for graphs with no vertex degree exceeding three. *Discrete Mathematics* **72**, 355–360.

S. Vadhan (2001). The Complexity of Counting in Sparse, Regular, and Planar Graphs. *SIAM Journal on Computing* **31**(2), 398–427.

L. G. Valiant (1979a). The complexity of computing the permanent. *Theoretical Computer Science* **8**, 189–201.

L. G. Valiant (1979b). The complexity of enumeration and reliability problems. *SIAM J. Computing* **8**(3), 410–421.

L. G. Valiant (2002). Expressiveness of matchgates. *Theoretical Computer Science* **289**(1), 457–471.

L. G. Valiant (2004). Holographic algorithms (extended abstract). In *Proc. 45th Annual IEEE Symposium on Foundations of Computer Science*, 17–19. Oct, Rome, Italy, IEEE Press, 306-315.

L. G. Valiant (2005). Completeness for parity problems. In *Proc. 11th International Computing and Combinatorics Conference, Aug, Kunming, China, LNCS, Vol. 3959, , 1-9*, 16–19.

L. G. Valiant (2006). Accidental algorithms. In *Proc. 47th Annual IEEE Symposium on Foundations of Computer Science*, 22–24. Oct, Berkeley, CA, IEEE Press, 509-517.

L. G. Valiant (2008). Holographic algorithms. *SIAM J. on Computing* **37**(5), 1565–1594. (Earlier version: *Electronic Colloquium on Computational Complexity*, Report TR-05-099, 2005).

M. Xia (2016). Base collapse of holographic algorithms. *STOC* 790–799.

M. Xia, P. Zhang & W. Zhao: (2007). Computational complexity of counting problems on 3-regular planar graphs. *Theor. Comput. Sci* **384**(1), 111–125.

Manuscript received 8 November 2016

Leslie G. Valiant
Harvard John A. Paulson School
of Engineering and Applied
Sciences,
Harvard University
Cambridge, MA 02138
valiant@seas.harvard.edu