

A UNIVERSALLY FASTEST ALGORITHM FOR MAX 2-SAT, MAX 2-CSP, AND EVERYTHING IN BETWEEN[†]

SERGE GASPERS AND GREGORY B. SORKIN

ABSTRACT. In this paper we introduce “hybrid” Max 2-CSP formulas consisting of “simple clauses”, namely conjunctions and disjunctions of pairs of variables, and general 2-variable clauses, which can be any integer-valued functions of pairs of boolean variables. This allows an algorithm to use both efficient reductions specific to AND and OR clauses, and other powerful reductions that require the general CSP setting. We use new reductions introduced here, and recent reductions such as “clause-learning” and “2-reductions” generalized to our setting’s mixture of simple and general clauses. We parametrize a hybrid instance by the fraction p of non-simple clauses.

We give an exact, exponential-time but polynomial-space algorithm that is the fastest known for $p = 0$, which includes the well-studied Max 2-Sat problem but also instances with arbitrary mixtures of AND and OR clauses; for an m -clause instance it runs in time $O^*(2^{m/6.321})$. The same algorithm is tied for fastest for general Max 2-CSP ($p = 1$), with running time $O^*(2^{m/5.263})$. The algorithm is the only one to treat mixtures of AND, OR, and general integer-valued clauses more efficiently than the general case, with intermediate running time bounds depending on the value of p . Since even a pure Max 2-Sat input instance may be transformed to a hybrid instance in the course of solving it, the algorithm’s efficiency and generality go hand in hand.

Our algorithm analysis and optimization use the familiar measure-and-conquer approach, but in a variation resulting in mathematical programs that are convex rather than quasi-convex, and can be solved efficiently and with a certificate of optimality. We produce a family of running-time upper-bound formulas, each optimized for instances with a particular value of p but valid for all instances.

CONTENTS

1. Introduction	2
2. Definitions	5
3. Algorithm and outline of the analysis	6
4. Some initial constraints	11
5. Simplification rules and their weight constraints	12
6. Some useful tools	16
7. Splitting reductions and preference order	20
8. Cubic instances	21
9. Instances of degree 4	25
10. Instances of degree 5	28
11. Instances of degree 6	31
12. Tuning the bounds	32
13. Conclusions	32
Acknowledgment	33
References	33
Appendix: Convex Program for Computing the Optimal Weights	35

Date: January 2011.

[†] The research was done largely during Serge Gaspers’ visit to IBM Research in July–September 2007, and Gregory Sorkin’s visit to the University of Bergen in April 2008, both with support from the Norwegian Research Council. Serge Gaspers was affiliated with the University of Bergen at that time and Gregory Sorkin with IBM Research. A preliminary version of this paper appeared in the proceedings of SODA 2009 [GS09].

1. INTRODUCTION

1.1. Treatment of “hybrid” Sat–CSP formulas. We show a polynomial-space algorithm that solves general instances of integer-valued Max 2-CSP (formally defined in Section 2), but that takes advantage of “simple” clauses, namely unit-weighted conjunctions and disjunctions, to reduce the running time. In a sense made precise near Remark 6, exclusive-or is the only boolean function we cannot treat efficiently.

Let us give a simple example. In the Max 2-CSP instance

$$(1) \quad (x_1 \vee x_2) + (x_2 \vee \overline{x_4}) + (x_2 \wedge x_3) + 3 \cdot (x_1 \vee x_3) + (2 \cdot (\overline{x_2}) - 5 \cdot x_4 + (x_2 \oplus x_4)),$$

the first two clauses are unit-weighted disjunctive clauses, the third clause is a unit-weighted conjunction, the fourth clause is a disjunction with weight 3, and the last clause is a general integer-valued CSP clause (any integer-valued 2-by-2 truth table). Thus this example has 3 simple clauses (the first three) and 2 non-simple clauses (the last two), for a fraction of non-simple clauses of $p = 2/5$. Both Max 2-Sat and Max 2-CSP have been extensively studied from the algorithmic point of view. For variable-exponential running times, the only two known algorithms faster than 2^n for Max 2-CSP (and Max 2-Sat) are those by Williams [Wil05] and Koivisto [Koi06], both with running time $O^*(2^{n/1.262})$. They employ beautiful ideas, but have exponential space complexity.

For clause-exponential running times, there has been a long series of improved algorithms; see Table 1. To solve Max 2-Sat, all early algorithms treated pure 2-Sat formulas. By using more powerful reductions closed over Max 2-CSP but not Max 2-Sat, the Max 2-CSP generalization of Scott and Sorkin [SS04] led to a faster algorithm. Then, several new Max 2-Sat specific reductions once again gave the edge to algorithms addressing Max 2-Sat instances particularly.

In this paper we get the best of both worlds by using reductions specific to Max 2-Sat (actually, we allow conjunctive as well as disjunctive clauses), but also using CSP reductions. While it is likely that Max 2-Sat algorithms will become still faster, we believe that further improvements will continue to use this method of combination.

1.2. Results. Let p be the fraction of non-simple clauses in the initial instance, no matter how this fraction changes during the execution of the algorithm. In example (1), $p = 2/5$. The algorithm we present here is the fastest known polynomial-space algorithm for $p = 0$ (including Max 2-Sat

Running Time	Problem	Space	Reference
$O^*(2^{m/2.879})$	Max 2-Sat	polynomial	Niedermeier and Rossmanith [NR00]
$O^*(2^{m/3.448})$	Max 2-Sat	polynomial	implicit by Bansal and Raman [BR99]
$O^*(2^{m/4})$	Max 2-Sat	polynomial	Hirsch [Hir00]
$O^*(2^{m/5})$	Max 2-Sat	polynomial	Gramm et al. [GHR03]
$O^*(2^{m/5})$	Max 2-CSP	polynomial	Scott and Sorkin [SS03]
$O^*(2^{m/5.263})$	Max 2-CSP	polynomial	Scott and Sorkin [SS04]
$O^*(2^{m/5.217})$	Max 2-Sat	polynomial	Kneis and Rossmanith [KR05]
$O^*(2^{m/5.769})$	Max 2-Sat	exponential	Kneis et al. [KMRR05]
$O^*(2^{m/5.5})$	Max 2-Sat	polynomial	Kojenikov and Kulikov [KK06]
$O^*(2^{m/5.769})$	Max 2-CSP	exponential	Scott and Sorkin [SS07a]
$O^*(2^{m/5.88})$	Max 2-Sat	polynomial	Kulikov and Kutzkov [KK07]
$O^*(2^{m/6.215})$	Max 2-Sat	polynomial	Raible and Fernau [RF08]
$O^*(2^{m/5.263})$	Max 2-CSP	polynomial	this paper
$O^*(2^{m/6.321})$	Max 2-Sat	polynomial	this paper

TABLE 1. A historical overview of algorithms for Max 2-Sat and Max 2-CSP

but also instances with arbitrary mixtures of AND and OR clauses); fastest for $0 < p < 0.29$ (where indeed no other algorithm is known, short of solving the instance as a case of general Max 2-CSP); and tied for fastest for $0.29 \leq p \leq 1$, notably for Max 2-CSP itself. For the well-known classes Max 2-Sat and Max 2-CSP, our algorithm has polynomial space complexity and running time $O^*(2^{m/6.321})$ and $O^*(2^{m/5.263})$, respectively.

For “cubic” instances, where each variable appears in at most three 2-variable clauses, our analysis gives running-time bounds that match and generalize the best known when $p = 0$ (including Max 2-Sat); improve on the best known when $0 < p < 1/2$; and match the best known for $1/2 \leq p \leq 1$ (including Max 2-CSP).

We derive running-time bounds that are optimized to the fraction p of non-simple clauses; see Table 2. Every such bound is valid for every formula, but the bound derived for one value of p may not be the best possible for a formula with a different value.

1.3. Method of analysis, and hybrid Sat–CSP formulas. Since a fair amount of machinery will have to be introduced before we can fully explain our analysis, let us first give a simplified overview of the method, including some new aspects of it in our application. Our algorithm reduces an instance to one or more smaller instances, which are solved recursively to yield a solution to the original instance. We view a Max 2-CSP instance as a constraint graph $G = (V, E \cup H)$ where vertices represent variables, the set of “light” edges E represents simple clauses and the set of “heavy” edges H represents general clauses. The reductions are usually local and change the constraint graph’s structure, and a related *measure*, in a predictable way.

For example, if G has two degree-4 vertices sharing two simple clauses, a “parallel-edge” reduction replaces the two simple clauses with one general clause, changing the vertices’ degrees from 4 to 3, giving a new constraint graph G' . With the measure μ including weights w_e and w_h for each simple and general clause (mnemonically, the subscripts refer to “edges” and “heavy” edges), and weights w_3 and w_4 for each vertex of degree 3 and 4, this reduction changes an instance’s measure by $\mu(G') - \mu(G) = -2w_e + w_h - 2w_4 + 2w_3$. An inductive proof of a running-time bound $O^*(2^{\mu(G)})$ will follow *if the measure change is non-positive*. Thus, we constrain that

$$-2w_e + w_h - 2w_4 + 2w_3 \leq 0.$$

An algorithm requires a set of reductions covering all instances: there must always be some applicable reduction. Just as above, each reduction imposes a constraint on the weights. One reduction’s constraint can weaken those of other reductions, by limiting the cases in which they are applied. For example, if we prioritize parallel-edge reduction, given as an example above (generalized to all degrees), we may assume that other reductions act on graphs without parallel edges. More usefully, “cut” reductions will allow us to assume that a graph has no small vertex cuts. Reductions like this producing a single instance, or any number of isomorphic instances, yield linear constraints (as in [SS03, SS04, SS07a]); reductions producing distinct instances yield nonlinear, convex constraints.

If a set of weights giving a measure μ satisfies all the constraints, the analysis results in a proof of a running-time bound $O^*(2^{\mu(G)})$ for an input instance G . To get the best possible running-time bound subject to the constraints, we wish to minimize $\mu(G)$. To avoid looking at the full degree spectrum of G , we constrain each vertex weight w_d to be non-positive, and then ignore these terms, resulting in a (possibly pessimistic) running-time bound $O^*(2^{|E|w_e + |H|w_h})$.

If G is a Max 2-Sat instance, to minimize the running-time bound is simply to minimize w_e subject to the constraints: as there are no heavy edges in the input instance, it makes no difference if w_h is large. This optimization will yield a small value of w_e and a large w_h . Symmetrically, if we are treating a general Max 2-CSP instance, where all edges are heavy, we need only minimize w_h . This optimization will yield weights w_e, w_h that are larger than the Max 2-Sat value of w_e but smaller than its w_h . For a hybrid instance with some edges of each type, minimizing $|E|w_e + |H|w_h$

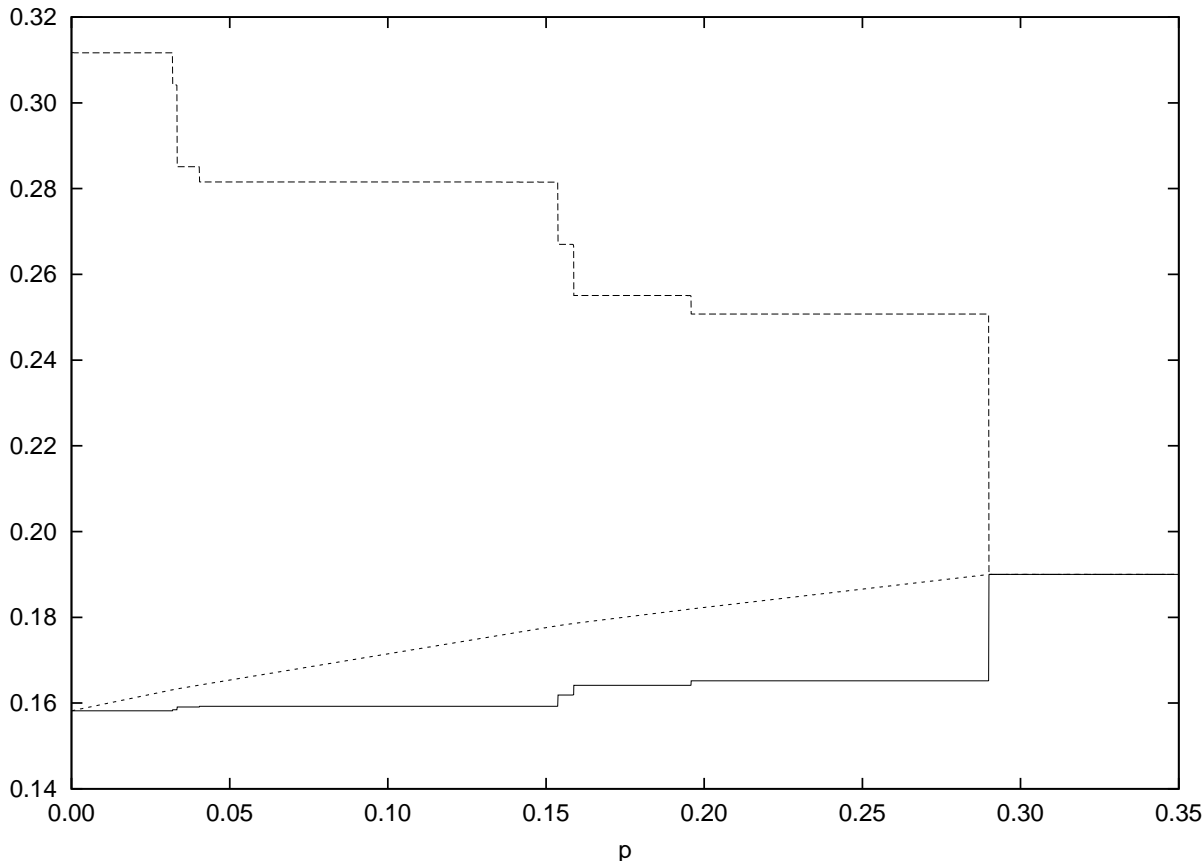


FIGURE 1. Plot of w_e (solid), w_h (dashed), and the running-time exponent $(1 - p)w_e + pw_h$ (dotted) versus the fraction p of non-simple 2-clauses. The three values are equal (and exactly 0.19) for $p > 0.29$. Both w_e and w_h appear to be piecewise constant: the resolution of the graph is in p increments of 0.0001, and all the small changes are meaningful.

is equivalent to minimizing $(1 - p)w_e + pw_h$, where $p = |H|/(|E| + |H|)$ is the fraction of non-simple clauses. This will result in weights w_e and w_h each lying between the extremes given by the pure 2-Sat and pure CSP cases; see Figure 1.

Thus, a new aspect of our approach is that it results in a family of nonlinear programs (NLPs), not just one: the NLPs differ in their objective functions, which are tuned to the fraction p of non-simple clauses in an input instance. The optimization done for a particular value of p , by construction, gives a running-time bound that is the best possible (within our methods) for an input instance with this fraction of non-simple clauses, but (because the constraints are the same in all the NLPs) that is valid for all instances; see the caption of Table 2.

1.4. Novel aspects of the analysis. Our introduction of the notion of hybrids between Max 2-Sat and Max 2-CSP, discussed above, is the main distinguishing feature of the present work. It yields a more general algorithm, applicable to CSP instances not just Sat instances, and gives better performance on Max 2-Sat by allowing both efficient Sat-specific reductions and powerful reductions that go outside that class. (The use of specific as well as general reductions distinguishes the hybrid approach from the ancient mathematical approach of working in a larger domain, and specifically from the work of Beigel and Eppstein for 3-coloring [BE05], and that of Scott and

Sorkin noted earlier.) This is surely not the final word on Max 2-Sat algorithms, but we expect new algorithms to take advantage of the hybrid approach.

A secondary point is that CSP reductions such as combining parallel edges or reducing on small cuts mean that in other cases it can be assumed that a graph has no parallel edges or small cuts. This potentially decreases the running-time bound (by weakening the corresponding constraints), and simplifies the case analysis, counter-balancing the complications of considering two types of edges.

Our analysis uses a now-common method, but with some novel aspects. Specifically, we analyze a reduction-based algorithm with a potential-function method akin to the measures used by [Kul97, Kul99], the quasi-convex analysis of [Epp06], the “measure and conquer” approach of [FGK09], the (dual to the) linear programming approach of [SS07a], and much older potential-function analyses in mathematics and physics. The goal is to solve a NLP giving a set of weights which minimizes a running-time bound, while respecting constraints imposed by the reductions. The hybrid view marks one change to this approach, since, as already discussed, it means that the objective function depends on the fraction of non-simple clauses, so there is a continuum of NLPs, not just one.

Our nonlinear programs are convex (those of [Epp06] are only quasi-convex), allowing them to be solved quickly and with certificates of optimality.

Also, it is common to make some assumptions about the weights, but we try to avoid this, instead only limiting the weights by the constraints necessitated by each reduction. This avoids unnecessary assumptions compromising optimality of the result, which is especially important in the hybrid realm where an assumption might be justified for Sat but not for CSP, or vice-versa. It also makes the analysis more transparent.

As is often the case with exact algorithms, regularity of an instance is important, and in our analysis we treat this with explicit weights penalizing regularity (motivated by a similar accounting for the number of 2-edges in a hypergraph in [Wah04], and the “forced moves” in [SS07a]). This introduces some extra bookkeeping but results in a more structured, more verifiable analysis.

We introduce several new reductions, including a 2-reduction combining ideas from [KK06] (for the Sat case) and [SS07a] (the CSP case), a “super 2-reduction”, and a generalization of the “clause-learning” from [KK07].

2. DEFINITIONS

We use the value 1 to indicate Boolean “true”, and 0 “false”. The canonical problem Max Sat is, given a boolean formula in conjunctive normal form (CNF), to find a boolean assignment to the variables of this formula satisfying a maximum number of clauses. Max 2-Sat is Max Sat restricted to instances in which each clause contains at most 2 literals.

We will consider a class more general than Max 2-Sat, namely integer-valued Max (2,2)-CSP; we will generally abbreviate this to Max 2-CSP. An instance (G, S) of Max 2-CSP is defined by a *constraint graph* (or multigraph) $G = (V, E)$ and a set S of *score functions*. There is a *dyadic* score function $s_e: \{0, 1\}^2 \rightarrow \mathbb{Z}$ for each edge $e \in E$, a *monadic* score function $s_v: \{0, 1\} \rightarrow \mathbb{Z}$ for each vertex $v \in V$, and (for bookkeeping convenience) a single *niladic* score “function” (really a constant) $s_\emptyset: \{0, 1\}^0 \rightarrow \mathbb{Z}$.

A candidate solution is a function $\phi: V \rightarrow \{0, 1\}$ assigning values to the vertices, and its score is

$$s(\phi) := \sum_{uv \in E} s_{uv}(\phi(u), \phi(v)) + \sum_{v \in V} s_v(\phi(v)) + s_\emptyset.$$

An optimal solution ϕ is one which maximizes $s(\phi)$.

The algorithm we present here solves any instance of Max 2-CSP with polynomial space usage, but runs faster for instances having a large proportion of “simple” clauses, namely conjunctions and disjunctions.

A hybrid instance $F = (V, E, H, S)$ is defined by its variables or vertices V , normal or *light* edges E representing conjunctive clauses and disjunctive clauses, *heavy* edges H representing arbitrary (integer-valued) clauses, and a set S of monadic functions and dyadic functions. Its light-and-heavy-edged constraint graph is $G = (V, E, H)$, though generally we will just think of the graph $(V, E \cup H)$; no confusion should arise. We will write $V(F)$ and $V(G)$ for the vertex set of an instance F or equivalently that of its constraint graph G .

In a graph G , we define the (*open*) *neighborhood* of a vertex u as $N(u) := \{v : uv \in E \cup H\} \setminus \{u\}$ (excluding u will not matter once we simplify our graphs and make them loopless), and the closed neighborhood as $N[u] := N(u) \cup \{u\}$. Generalizing, a *set* of vertices, U , has (open) neighborhood $N(U) = (\bigcup_{u \in U} N(u)) \setminus U$, and (open) *second neighborhood* $N^2(U) = N(N(U)) \setminus U$. For a single vertex u , define $N^2(u) := N^2(\{u\})$. By definition, U , $N(U)$, and $N^2(U)$ are disjoint.

We define the degree $\deg(u)$ of a vertex u to be the number of edges incident on u where loops are counted twice, and the *degree* (or *maximum degree*) of a formula F (or its constraint graph G) to be the maximum of its vertex degrees. Without loss of generality we will assume that there is at most one score function for each vertex, though we will allow multiple edges. Then, up to constant factors the space required to specify an instance F with constraint graph $G = (V, E, H)$ is the *instance size*

$$(2) \quad |F| = 1 + |V| + |E| + |H|.$$

We use the symbol \square to end the description of a reduction rule or the analysis of a case, and \square to end a proof.

3. ALGORITHM AND OUTLINE OF THE ANALYSIS

We will show an algorithm (sketched as Algorithm 1) which, on input of a hybrid instance F , returns an optimal coloring ϕ of F 's vertices in time $O^*(2^{w_e|E|+w_h|H|})$, which is to say in time

$$(3) \quad T(F) \leq \text{poly}(|F|)2^{w_e|E|+w_h|H|}$$

for some polynomial $\text{poly}(\cdot)$.

3.1. Algorithm and general arguments. The algorithm is recursive: on input of an instance F , in time polynomial in the instance size $|F|$, F is *reduced* to a single instance F' (a *simplification*) or to several instances F_1, \dots, F_k (a *splitting*), each of smaller size; the algorithm solves the reduced instance(s) recursively; and, again in time $\text{poly}(|F|)$, the algorithm constructs an optimal solution to F from the solutions of the reduced instances.

The central argument (corresponding to the analysis for line 13 of Algorithm 1) is to establish (3) for *simplified* formulas of maximum degree at most 6. We do this shortly, in Lemma 1, with the bulk of the paper devoted to verifying the lemma's hypotheses.

Given Lemma 1, we then establish a similar running-time bound for instances F of degree at most 6 which are not simplified, that is, instances to which we may apply one or more of the simplifications of Procedure 2 (the analysis referred to by line 8 in Algorithm 1), and for instances of arbitrary degree (the argument alluded to in line 5 of Algorithm 1).

3.2. Central argument. The main argument is to establish (3) for simplified formulas of maximum degree at most 6. We will prove that

$$(4) \quad T(F) = O(|F|^k 2^{\mu(F)}),$$

which suffices if (as we will ensure) for some constant C and every simplified instance F of degree at most 6, the *measure* $\mu(F)$ satisfies

$$(5) \quad \mu(F) \leq w_e|E| + w_h|H| + C.$$

Algorithm 1 Outline of algorithm and analysis

- 1: **Input:** A hybrid Max 2-Sat / 2-CSP instance F .
 - 2: **Output:** An optimal coloring ϕ of the vertices of F .
 - 3: **if** F has any vertex v of degree ≥ 7 **then**
 - 4: Split on $\phi(v) = 0$ and $\phi(v) = 1$ to obtain F_1, F_2 , recursively solve the instances F_1 and F_2 and return the best assignment for F .
 - 5: (Analysis: Inductively establish running time, using that both F_1 and F_2 have at least 7 edges fewer than F .)
 - 6: **end if**
 - 7: Simplify F . (See Procedure 2.)
 - 8: (Analysis: Establish running-time bound for general instances, using a bound for simplified instances.)
 - 9: **if** F is nonempty **then**
 - 10: Apply first applicable splitting reduction, obtaining F_1, \dots, F_k .
 - 11: Simplify each of F_1, \dots, F_k .
 - 12: Recursively solve F_1, \dots, F_k and return the best assignment for F .
 - 13: (Analysis: Inductively establish running-time bound for simplified instances of maximum degree ≤ 6 , using $\sum_{i=1}^k 2^{\mu(F_i)} \leq 2^{\mu(F)}$.)
 - 14: **end if**
-

Procedure 2 Simplification procedure

- 1: **Input:** A hybrid instance F
 - 2: **while** Any of the following simplification rules is applicable **do**
 - 3: Apply the first applicable simplification: combine parallel edges; remove loops; 0-reduction; delete a small component; delete a decomposable edge; half-edge reduction; 1-reduction; 1-cut; 2-reduction; 2-cut.
 - 4: **end while**
 - 5: Return the resulting simplified instance.
-

In the following lemma's application, the class \mathcal{F} will consist of simplified hybrid formulas of degree at most 6.

Lemma 1 (Main Lemma). *For a family \mathcal{F} of formulas, suppose there exists an algorithm A and a constant $c \geq 1$, such that on input of any instance $F \in \mathcal{F}$, A either solves F directly in time $O(1)$, or decomposes F into instances $F_1, \dots, F_k \in \mathcal{F}$, solves these recursively, and inverts their solutions to solve F , using time $O(|F|^c)$ for the decomposition and inversion (but not the recursive solves). Further suppose that for a given measure μ ,*

$$(6) \quad (\forall F \in \mathcal{F}) \quad \mu(F) \geq 0,$$

and, for any decomposition done by algorithm A ,

$$(7) \quad (\forall i) \quad |F_i| \leq |F| - 1, \text{ and}$$

$$(8) \quad 2^{\mu(F_1)} + \dots + 2^{\mu(F_k)} \leq 2^{\mu(F)}.$$

Then A solves any instance $F \in \mathcal{F}$ in time $O(|F|^{c+1})2^{\mu(F)}$.

We will often work with the equivalent to (8), that

$$(8') \quad \sum_{i=1}^k 2^{\mu(F_i) - \mu(F)} \leq 1.$$

Proof. The result follows easily by induction on $|F|$. Without loss of generality, we may replace the hypotheses' O statements with simple inequalities (substitute a sufficiently large leading constant, which then appears everywhere and has no relevance), and likewise for the base case assume that we can return the solution \emptyset to an empty instance in time 1. If an instance F is solved in time 1, then $T(F) = 1 \leq |F|^{c+1} 2^{\mu(F)}$. Otherwise, where $T(F)$ denotes the time taken to solve an instance,

$$\begin{aligned} T(F) &\leq |F|^c + \sum_{j=1}^k T(F_j) && \text{(by hypothesis)} \\ &\leq |F|^c + \sum |F_i|^{c+1} 2^{\mu(F_i)} && \text{(by the inductive hypothesis)} \\ &\leq |F|^c + (|F| - 1)^{c+1} \sum 2^{\mu(F_i)} && \text{(by hypothesis (7))} \\ &\leq |F|^c + (|F| - 1)^{c+1} 2^{\mu(F)} && \text{(by hypothesis (8))} \\ &\leq |F|^{c+1} 2^{\mu(F)}. \end{aligned}$$

The final inequality uses that $\mu(F) \geq 0$ and holds for any $c \geq 1$. \square

The main work of the paper will be to find a set of decompositions and a measure μ such that the decompositions satisfy inequality (7), μ satisfies inequality (6), and (more interestingly) μ satisfies inequality (5) for some small values of w_e and w_h , and finally, for every decomposition, μ satisfies inequality (8).

3.3. Measure. For an instance F of (maximum) degree at most 6, we define a measure $\mu(F)$ as a sum of weights associated with light edges, heavy edges, and vertices of various degrees (at most 6), and constants associated with the maximum degree d of F and whether F is regular (for all the degree criteria treating light and heavy edges alike):

$$(9) \quad \mu(F) := \nu(F) + \delta(F), \text{ with}$$

$$(10) \quad \nu(F) := |E|w_e + |H|w_h + \sum_{v \in V} w_{\deg(v)},$$

$$(11) \quad \delta(F) := \sum_{d=4}^6 \chi(\max \deg(G) \geq d) C_d + \sum_{d=4}^6 \chi(G \text{ is } d\text{-regular}) R_d.$$

Here $\chi(\cdot)$ is the indicator function: 1 if its argument is true, 0 otherwise.

To satisfy condition (5) it is sufficient that

$$(12) \quad (\forall d) \quad w_d \leq 0;$$

this is also necessary for large regular instances. Since we are now only considering instances of degree ≤ 6 , we interpret “ $\forall d$ ” to mean for all $d \in \{0, 1, \dots, 6\}$.

3.4. Peripheral arguments. We first dispense with non-simplified instances.

Lemma 2. *Suppose that every simplified Max 2-CSP instance F of degree at most $D \leq 6$ can be solved in time $\text{poly}_1(|F|) 2^{\mu(F)}$. Suppose also that*

- (1) *simplifying F (or determining that F is already simplified) takes time at most $\text{poly}_2(|F|)$,*
- (2) *any instance F' obtained from simplifying F satisfies $|F'| \leq |F| - 1$ and $\mu(F') \leq \mu(F) + C'$ for some positive constant C' , and*
- (3) *the simplification can be reversed in time at most $\text{poly}_2(|F|)$ to recover an optimal solution to F from any optimal solution of F' .*

Then any instance F of degree at most D can be solved in time $\text{poly}(|F|)2^{\mu(F)}$, with $\text{poly}(x) := \text{poly}_2(x) + 2^{C'} \text{poly}_1(x)$.

Proof. Since simplifying reduces the instance size, a solution to the original instance F can be obtained in time

$$\begin{aligned} T(F) &\leq \text{poly}_2(|F|) + T(F') \\ &\leq \text{poly}_2(|F|) + \text{poly}_1(|F'|)2^{\mu(F')} \\ &\leq \text{poly}_2(|F|) + \text{poly}_1(|F|)2^{\mu(F)+C'} \\ &\leq (\text{poly}_2(|F|) + 2^{C'} \text{poly}_1(|F|))2^{\mu(F)} \\ &= \text{poly}(|F|)2^{\mu(F)}. \end{aligned}$$

□

The lemma's hypotheses (1) and (3) will be satisfied by construction. Hypothesis (2) is assured if we constrain that, for each simplification rule taking F to F' ,

$$(13) \quad \nu(F') \leq \nu(F),$$

since by transitivity the same inequality then holds for any sequence of simplifications starting with F and ending with a simplified instance F' , and the desired inequality $\mu(F') = \nu(F) + \delta(F) - \delta(F') \leq \nu(F) + C'$ follows by boundedness of δ and choosing C' sufficiently large.

Finally, we dispense with instances of high degree, the argument alluded to in line 5 of Algorithm 1.

Lemma 3. *Suppose that every Max 2-CSP instance F of degree at most 6 can be solved in time $O(|F|^{k_1} 2^{w_e|E|+w_h|H|})$, with $w_e, w_h \geq 1/7$. Then for some sufficiently large k , every instance F can be solved in time $O(|F|^k 2^{w_e|E|+w_h|H|})$.*

Proof. As in the proof of Lemma 1, without loss of generality we may replace the O statement in the hypothesis with a simple inequality. If F has any vertex v of degree at least 7, we will set $\phi(v)$ to 0 and 1 to generate instances F_0 and F_1 respectively, solve them recursively, and note that the solution to F is that of the better of F_0 and F_1 , extended with the corresponding value for $\phi(v)$. We may assume that the splitting and its reversal together take time at most $|F|^{k_2}$.

Ensure that $k \geq k_1$ is large enough that for all $x \geq 2$, $x^{k_2} \leq x^k - (x-1)^k$, and note that the hypothesis remains true replacing k_1 with k .

The proof is by induction on F . If F has no vertex of degree at least 7 then we are already done. Otherwise reduce F to F_1 and F_2 , each having at least 7 fewer (light and/or heavy) edges than F . By induction we may assume the bound for $T(F_1)$ and $T(F_2)$, so

$$\begin{aligned} T(F) &\leq |F|^{k_2} + 2(|F| - 1)^k 2^{w_e|E|+w_h|H|-7 \cdot 1/7} \\ &= |F|^{k_2} + (|F| - 1)^k 2^{w_e|E|+w_h|H|} \\ &\leq |F|^k 2^{w_e|E|+w_h|H|}. \end{aligned}$$

The worst case for the last inequality is when $w_e|E| + w_h|H| = 0$ (it is nonnegative), and in that case the inequality follows by the construction of k . □

3.5. Optimizing the measure. The task of the rest of the paper is to produce the comprehensive set of reductions hypothesized by Lemma 1 (to any formula there should be some reduction we can apply) and a measure μ , satisfying the hypotheses, with w_e as small as possible. (More generally, if there are $m(1-p)$ conjunctions and mp general integer-valued clauses, we wish to minimize

p $\Delta(F)$	0			0.05			0.1		
	w_e	w_h	w	w_e	w_h	w	w_e	w_h	w
3	0.10209	0.23127	0.10209	0.10209	0.23125	0.10855	0.10209	0.23125	0.11501
4	0.14662	0.31270	0.14662	0.14662	0.31270	0.15493	0.15023	0.26951	0.16216
5	0.15518	0.30728	0.15518	0.15637	0.27997	0.16255	0.15640	0.27951	0.16871
≥ 6	0.15819	0.31029	0.15819	0.15912	0.28223	0.16527	0.15912	0.28223	0.17143

p $\Delta(F)$	0.2			0.3			1		
	w_e	w_h	w	w_e	w_h	w	w_e	w_h	w
3	0.10209	0.23125	0.12793	0.10209	0.23125	0.14084	0.16667	0.16667	0.16667
4	0.15023	0.26951	0.17409	0.15023	0.26951	0.18601	0.18750	0.18750	0.18750
5	0.15640	0.27951	0.18102	0.19000	0.19000	0.19000	0.19000	0.19000	0.19000
≥ 6	0.16520	0.25074	0.18231	0.19000	0.19000	0.19000	0.19000	0.19000	0.19000

TABLE 2. Values of w_e , w_h and $w := pw_h + (1-p)w_e$ according to the fraction p of heavy edges and the maximum degree $\Delta(F)$ of a formula F . For any pair (w_e, w_h) in the table, a running-time bound of $O^*(2^{m \cdot ((1-p)w_e + pw_h)})$ is valid for every formula, regardless of its fraction $p(F)$ of non-simple clauses, but the pair obtained when the table's p equals $p(F)$ gives the best bound.

$m(1-p)w_e + mpw_h$ or equivalently $(1-p)w_e + pw_h$, but for the discussion here we will just think in terms of minimizing w_e .)

For each reduction, the hypothesized constraint (7) will be trivially satisfied, and it will be straightforward to write down a constraint ensuring (8'). We then solve the nonlinear program of minimizing w_e subject to all the constraints.

Minimizing w_e for a given set of constraints can be done with an off-the-shelf nonlinear solver (see Section 8.6), but finding a set of reductions resulting in a small value of w_e remains an art. It consists of trying some set of reductions, seeing which ones' constraints are tight in an optimal solution, and trying to replace these reductions with more favorable ones.

With the constraints established in the next sections, we will obtain our main result.

Theorem 4. *Let F be an instance of integer-weighted Max 2-CSP in which each variable appears in at most $\Delta(F)$ 2-clauses, and there are $(1-p(F))m$ conjunctive and disjunctive 2-clauses, and $p(F)m$ other 2-clauses. Then, for any pair of values w_e, w_h in Table 2 (not necessarily with the table's p equal to $p(F)$), the above algorithm solves F in time $O^*(2^{m \cdot ((1-p(F))w_e + p(F)w_h)})$. When the table's $p = p(F)$, we obtain our best bound, $O^*(2^{m \cdot ((1-p)w_e + pw_h)}) = O^*(2^{mw})$.*

Proof. Corollary of Lemma 1, solving the mathematical program given by the various constraints given in the next sections and minimizing $pw_h + (1-p)w_e$. \square

Which of the constraints are tight strongly depends on p and $\Delta(F)$.

3.6. The measure's form. Let us explain the rather strange form of the measure. Ideally, it would be defined simply as ν , and indeed for the measure we ultimately derive, all of our simplifications and most of our splittings satisfy the key inequality (8') with ν alone in place of μ . Unfortunately, for regular instances of degrees 4, 5, and 6, satisfying this constraint would require a larger value of w_e . Viewing (8') equivalently as

$$\sum_{i=1}^k 2^{\mu(F_i) - \mu(F)} \leq 1,$$

adding a cost R_d to the measure of a d -regular instance F means that if a d -regular instance F is reduced to nonregular instances F_1 and F_2 of degree d , each difference $\mu(F_i) - \mu(F)$ is smaller by R_d than the corresponding difference $\nu(F_i) - \nu(F)$. We will therefore want

$$(14) \quad (\forall d \in \{4, 5, 6\}) \quad R_d \geq 0.$$

Of course, if a nonregular instance F of degree d is reduced to instances F_i of degree d one or more of which is regular, there will be a corresponding penalty: for each d -regular F_i , $\mu(F_i) - \mu(F)$ is $\nu(F_i) - \nu(F) + R_d$.

Indeed, for each splitting reduction we will have to consider several cases. Typically, the “baseline” case will be the reduction of a nonregular instance to two nonregular instances. In this case μ and ν are equivalent, and if we know for example that $\nu(F_i) - \nu(F) \leq x_i$, our nonlinear program constrains that $2^{x_1} + 2^{x_2} \leq 1$.

If we reduce starting from a regular instance, the nature of the reductions is such that, generically, we will get less favorable bounds $\nu(F_i) - \nu(F) \leq x'_i$ (the values x'_i will be larger than the x_i were), but we also get a “reward” (a further decrease of R_d) for whichever of F_1 and F_2 are not also regular. If we reduce starting from a nonregular instance but producing one or more regular children, we will consider various possibilities.

The case where a nonregular instance of degree d produced a regular instance F_i of degree $< d$, can be dispensed with simply by choosing C_d sufficiently large, to reap whatever additional reward is needed. Our splitting rules are generally local and will never increase measure by more than a constant, so some constant C_d suffices. Also, our reductions never increase the degree of an instance (each F_i has degree at most that of F), so C_d will never work against us, and there is no harm in choosing it as large as we like. Thus, we never need to consider the particulars of cases where the instance degree decreases, nor the values C_d .

The remaining cases where a nonregular instance has regular children will be considered on a case-by-case basis for each reduction. Generally, for a child to become regular means that, beyond the constraint-graph changes taken into account in the baseline case (with the child nonregular), some additional vertices (those of degree less than d) must have been removed from the instance by simplifications. Accounting for these implies a further decrease in measure that compensates for the increase by R_d .

4. SOME INITIAL CONSTRAINTS

We have already derived one constraint for μ , namely (12), and we will now introduce some notation and derive several more constraints.

Let us write $w(v)$ for the weight of a vertex v (so $w(v) = w_d$ for a vertex of degree d), and similarly $w(e)$ for the weight of an edge (w_e or w_h depending on whether e is light or heavy). Sometimes it will be helpful to think of $\nu(F)$ as

$$(15) \quad \nu(F) = \sum_{v \in V} \left(w(v) + \frac{1}{2} \sum_{e: v \in e} w(e) \right),$$

the sum of the weights of the vertices and their incident *half-edges*. For convenience, we define (and thus constrain)

$$(16) \quad a_d = w_d + \frac{1}{2} d w_e.$$

Thus, a_d is equal to the summand in (15) for a vertex all of whose incident edges are light, and smaller otherwise.

We require $\mu(F) \geq 0$ for all instances. Considering regular Max 2-Sat instances with degree d ($d = 0, \dots, 6$), this implies that

$$(17) \quad (\forall d) \quad a_d \geq 0.$$

(For $d \leq 3$, (17) is implied by $\delta(F) = 0$, with (15) and (16). For $d \geq 4$, positivity of ν might give positive measure to K_d even if $\delta(K_d)$ were negative, but then a graph consisting of sufficiently many copies of K_d would still have negative measure.) If we also constrain that

$$(18) \quad (\forall d \in \{4, 5, 6\}) \quad C_d, R_d \geq 0,$$

then we have assured that $\mu(F) \geq 0$ for all instances. In the end, constraint (18) will not be tight and so there is no loss in making the assumption.

Were it the case that $w_h \leq w_e$, then we could simply transform each light edge into a heavy one, reducing the measure, and getting a better time bound for solving an instance of Max 2-CSP than an instance of Max 2-Sat or a hybrid instance. Thus if we are to gain any advantage from considering Max 2-Sat or hybrid instances, it must be that

$$(19) \quad w_e \leq w_h.$$

In the end we will find that this constraint is not tight, and so there is no cost to making the assumption.¹

For intuitive purposes let us leap ahead and mention that we will find that $a_0 = a_1 = a_2 = 0$, (thus $w_0 = 0$, $w_1 = -\frac{1}{2}w_e$, and $w_2 = -w_e$), while $0 < a_3 < \dots < a_6$. Per (19) above, $w_h \geq w_e$. Typically we will find that $w_h \leq 2w_e$, but not always. (Even where this fails to hold, notably for cubic Max 2-Sat, we can still replace two conjunctions or disjunctions on the same variables with one CSP edge: decreasing the degrees of the incident vertices decreases the measure enough to make up for the increase of $w_h - 2w_e$.) This “intuition” has changed several times as the paper has evolved, which supports the value of making as few assumptions as possible, instead just writing down constraints implied by the reductions.

5. SIMPLIFICATION RULES AND THEIR WEIGHT CONSTRAINTS

We use a number of simplification rules (reductions of F to a single simpler instance F_1 or F'). Some of the simplification rules are standard, the CSP 1-reductions are taken from [SS07a], the CSP 2-reductions combine ideas from [SS07a] and [KK06], and a “super 2-reduction” is introduced here. For vertices of degree 5 we use a splitting reduction taken from [KK07] that we generalize to hybrid instances.

We have already ensured constraint (6) by (17) and (18), so our focus is on ensuring that each reduction satisfies (8'). Since each splitting is followed by an (unpredictable) sequence of simplifications, to have any hope of satisfying (8') it is essential that each simplification from any F to F' satisfies

$$(20) \quad \nu(F') \leq \nu(F);$$

in any case this inequality is required by Lemma 2 (it duplicates inequality (13)). Constraint (7) of Lemma 1 will be trivially satisfied by all our simplifications and splittings.

Recapitulating, in this section we show that (20) is satisfied by all our simplifications. Ensuring (8') will come when we look at the splitting rules, and the measure component δ we are ignoring here.

5.1. Combine parallel edges. Two parallel edges (light or heavy) with endpoints x and y may be collapsed into a single heavy edge. This means that the “transformed” instance F' (F_1 in Lemma 1, with $k = 1$) is identical to F except that the two score functions $s_{xy}(\phi(x), \phi(y))$ and $s'_{xy}(\phi(x), \phi(y))$ in F are replaced by their sum $s''_{xy}(\phi(x), \phi(y))$ in F' . If one of the endpoints, say x , of the two

¹For the most part we will only write down constraints that are *necessary*, typically being required for some reduction to satisfy (8'), but we make a few exceptions early on.

parallel edges has degree 2, collapse the parallel edges and immediately apply a 1-reduction (see 5.7) on x (of degree 1), which removes x from the constraint graph. To ensure (20) we constrain

$$(21) \quad (\forall d \geq 2) \quad -a_2 - a_d + a_{d-2} \leq 0 :$$

the left-hand side is $\nu(F') - \nu(F)$ thought of as the subtraction of a vertex of degree 2, a vertex of degree d and the addition of a vertex of degree $d - 2$. For the case that x and y have degree $d \geq 3$, we constrain

$$(22) \quad (\forall d \geq 3) \quad -2a_d + 2a_{d-1} - w_e + w_h \leq 0 :$$

the left-hand side is $\nu(F') - \nu(F)$ thought of as replacing two vertices of degree d by two vertices of degree $d - 1$ and replacing a light edge by a heavy edge. (Remember that the degree of a vertex is the number of incident edges rather than the number of distinct neighbors.) If $\deg(x) \neq \deg(y)$, the resulting constraint is a half-half mixture of a constraint (22) with $d = \deg(x)$ and another with $d = \deg(y)$, and is thus redundant.

By construction, the score functions of F' and F are identical, so an optimal solution ϕ' for F' is an optimal solution ϕ of F (no transformation is needed). \square

Applying this reduction whenever possible, we may assume that the instance has no parallel edges.

Note that we cannot hope to combine simple clauses (conjunctions and disjunctions) and still take advantage of their being simple clauses rather than general CSP clauses: $(x \vee y) + (\bar{x} \vee \bar{y}) = 1 + (x \oplus y)$, the additive 1 is irrelevant, and the XOR function is not simple.

5.2. Remove loops. If the instance includes any edge $xx \in E \cup H$, the nominally dyadic score function $s_{xx}(\phi(x), \phi(x))$ may be replaced by a (or incorporated into an existing) monadic score function $s_x(\phi(x))$. This imposes the constraints

$$(23) \quad (\forall d \geq 2) \quad -a_d + a_{d-2} \leq 0.$$

\square

As this constraint is stronger than (21), we may ignore constraint (21).

With this and the edge-combining reduction, we may at all times assume the constraint graph is simple.

5.3. Delete a vertex of degree 0 (0-reduction). If v is a vertex of degree 0, reduce the instance F to F' by deleting v and its monadic score function s_v , solve F' , and obtain an optimal solution of F by augmenting the solution of F' with whichever coloring $\phi(v)$ of v gives a larger value of $s_v(\phi(v))$. Constraint (7) is satisfied, since $|F'| = |F| - 1$. Constraint (20) is satisfied if and only if $-w_0 \leq 0$. On the other hand, for a useful result we need each $w_d \leq 0$ (inequality (12)), implying that $w_0 = 0$, and thus

$$(24) \quad a_0 = 0.$$

We will henceforth ignore vertices of degree 0 completely. \square

5.4. Delete a small component. For a constant C (whose value we will fix in the splitting reduction (reduction 7.1)), if the constraint graph G of F has components G' and G'' with $1 \leq |V(G'')| < C$ ($|V(G')|$ is arbitrary), then F may be reduced to F' with constraint graph G' . The reduction and its correctness are obvious, noting that F'' may be solved in constant time. Since $\nu(F') - \nu(F) \leq -\sum_{v \in V(G)} a_{\deg(v)}$, it is immediate from (17) that (20) is satisfied. \square

5.5. Delete a decomposable edge. If a dyadic score function $s_{xy}(\phi(x), \phi(y))$ can be expressed as a sum of monadic scores, $s'_x(\phi(x)) + s'_y(\phi(y))$, then delete the edge and add s'_x to the original s_x , and s'_y to s_y . For example, $(x_1 \oplus x_2) + (x_1 \wedge \overline{x_2}) + 3 \cdot (x_1 \wedge x_2)$ may be decomposed as $2 \cdot (x_1) + (x_2)$. If x and y have equal degrees, the constraint imposed is that $(\forall d \geq 1) -w_e - 2w_d + 2w_{d-1} \leq 0$, or equivalently,

$$(25) \quad (\forall d \geq 1) \quad -a_d + a_{d-1} \leq 0.$$

(The $d = 1$ case was already implied by (24) and (17).) As in (22), inequalities for degree pairs are a mixture of those for single degrees. Note that we may ignore constraint (23) now as it is weaker than (25). \square

Three remarks. First, together with (24), (25) means that

$$(26) \quad 0 = a_0 \leq a_1 \leq \dots \leq a_6.$$

Second, if an edge is not decomposable, the assignment of either endpoint has a (nonzero) bearing on the optimal assignment of the other, as we make precise in Remark 5. We will exploit this in Lemma 7, which shows how “super 2-reduction” opportunities (reduction 6.1) are created.

Remark 5. *Let*

$$\text{bias}_y(i) := s_{xy}(i, 1) - s_{xy}(i, 0),$$

the “preference” of the edge function s_{xy} for setting $\phi(y) = 1$ over $\phi(y) = 0$ when x is assigned $\phi(x) = i$. Then s_{xy} is decomposable if and only if $\text{bias}_y(0) = \text{bias}_y(1)$.

Proof. We have that s_{xy} is decomposable if and only if its 2-by-2 table of function values has rank 1, which is equivalent to equality of the two diagonal sums, $s_{xy}(0, 1) + s_{xy}(1, 0) = s_{xy}(0, 0) + s_{xy}(1, 1)$, which in turn is equivalent to $s_{xy}(0, 1) - s_{xy}(0, 0) = s_{xy}(1, 1) - s_{xy}(1, 0)$, i.e., $\text{bias}_y(0) = \text{bias}_y(1)$. \square

Finally, when some vertices and their incident edges are deleted from a graph, we may think of this as the deletion of each vertex and its incident half-edges (which typically we will account for explicitly) followed (which we may not account for) by the deletion of any remaining half-edges and the concomitant decrease in the degrees of their incident vertices (for edges one of whose endpoints was deleted and one not). A “half-edge deletion” and vertex degree decrease is precisely what is characterized by the left-hand side of (25), so it cannot increase the measure ν . Even though such simplifications take place on an intermediate structure that is more general than a graph, and that we will not formalize, for convenient reference we will call this a half-edge reduction.

5.6. Half-edge reduction. Delete a half-edge, and decrease the degree of its incident vertex. By (25), this does not increase the measure.

5.7. Delete a vertex of degree 1 (1-reduction). This reduction comes from [SS07a], and works regardless of the weight of the incident edge. Let y be a vertex of degree 1, with neighbor x . Roughly, we use the fact that the optimal assignment of y is some easily-computable function of the assignment of x , and thus y and its attendant score functions $s_y(\phi(y))$ and $s_{xy}(\phi(x), \phi(y))$ can be incorporated into $s_x(\phi(x))$.

We take a precise formulation from [SS07a]. Here V is the vertex set of F , E is the set of all edges (light and heavy), and S is the set of score functions.

Reducing (V, E, S) on y results in a new instance (V', E', S') with $V' = V \setminus y$ and $E' = E \setminus xy$. S' is the restriction of S to V' and E' , except that for all “colors” $C \in \{0, 1\}$ we set

$$s'_x(C) = s_x(C) + \max_{D \in \{0, 1\}} \{s_{xy}(CD) + s_y(D)\}.$$

Note that any coloring ϕ' of V' can be extended to a coloring ϕ of V in two ways, depending on the color assigned to y . Writing (ϕ', D) for the extension in which $\phi(y) = D$, the defining property

of the reduction is that $S'(\phi') = \max_D S(\phi', D)$. In particular, $\max_{\phi'} S'(\phi') = \max_{\phi} S(\phi)$, and an optimal coloring ϕ' for the instance (V', E', S') can be extended to an optimal coloring ϕ for (V, E, S) . This establishes the validity of the reduction.

Since the reduction deletes the vertex of degree 1 and its incident edge (light, in the worst case), and decreases the degree of the adjacent vertex, to ensure (20), we constrain that $(\forall d \geq 1) -w_1 - w_e - w_d + w_{d-1} \leq 0$, or equivalently that

$$(\forall d \geq 1) \quad a_{d-1} - a_d - a_1 \leq 0,$$

which is already ensured by (26). \square

5.8. 1-cut. Let x be a cut vertex isolating a set of vertices A , $2 \leq |A| \leq 10$. (The 1-cut reduction extends the 1-reduction, thought of as the case $|A| = 1$.) Informally, for each of $\phi(x) = 0, 1$ we may determine the optimal assignments of the vertices in A and the corresponding optimal score; adding this score function to the original monadic score s_x gives an equivalent instance F' on variables $V \setminus A$. With A of bounded size, construction of F' , and extension of an optimal solution of F' to one of F , can be done in polynomial time. (Formal treatment of a more general “cut reduction” on more general “Polynomial CSPs” can be found in [SS07b].)

This simplification imposes no new constraint on the weights. Vertices in A and their incident half-edges are deleted, and any remaining half-edges (those incident on x) are removed by half-edge reductions (reduction 5.6); by (26), neither increases the measure ν . \square

5.9. Contract a vertex of degree 2 (2-reduction). Let y be a vertex of degree 2 with neighbors x and z . Then y may be contracted out of the instance: the old edges xy , yz , and (if any) xz are replaced by a single new edge xz which in general is heavy, but is light if there was no existing edge xz and at least one of xy and yz was light.

The basics are simple, but care is needed both because of the distinction between light and heavy edges and because we insist that the constraint graph be simple, and the 2-reduction is the one operation that has the capacity to (temporarily) create parallel edges and in the process change the vertex degrees. We consider two cases: there is an edge xz ; and there is no edge xz .

If there is an edge xz then x and z both have degree 3 or more by Simplification 5.8, we use the general Max 2-CSP 2-reduction from [SS07a]. Arguing as in the 1-reduction above, here the optimal assignment of y depends only on the assignments of x and z , and thus we may incorporate all the score terms involving y , namely $s_y(\phi(y)) + s_{xy}(\phi(x), \phi(y)) + s_{yz}(\phi(y), \phi(z))$, into a new $s'_{xz}(\phi(x), \phi(z))$, which is then combined with the original $s_{xz}(\phi(x), \phi(z))$. The effect is that y is deleted, three edges (in the worst case all light) are replaced by one heavy edge, and the degrees of x and z decrease by one. If $\deg(x) = \deg(y) = d$, $\nu(F') - \nu(F) \leq 0$ is assured by $-w_2 - 3w_e + w_h - 2w_d + 2w_{d-1} \leq 0$, or equivalently

$$(\forall d \geq 3) \quad -a_2 - w_e + w_h - 2a_d + 2a_{d-1} \leq 0,$$

which is already ensured by (17) and (22). As in (25), inequalities for pairs $\deg(x) \neq \deg(y)$ are a mixture of those for single degrees. If xy or yz is heavy, then $\nu(F') - \nu(F) \leq -w_h + w_e$, and we will capitalize on this later.

Finally, we consider the case where there was no edge xz . If xy and yz are both heavy, then as in the first case we apply the general Max 2-CSP reduction to replace them with a heavy edge xz , giving $\nu(F') - \nu(F) \leq -2w_h + w_h - w_2 = -a_2 - w_h + w_e \leq -w_h + w_e$.

Otherwise, at least one of xy and yz is light, and we show that the resulting edge xz is light. (For pure Sat formulas, this is the “frequently meeting variables” rule of [KK06].) Without loss of generality we assume that xy is the conjunctive constraint $x \vee y$ or the disjunction $x \wedge y$ (what is relevant is that the clause’s score is restricted to $\{0, 1\}$, and is monotone in $\phi(y)$). We define a *bias*

$$(27) \quad \text{bias}_y(i) = [s_y(1) - s_y(0)] + [s_{yz}(1, i) - s_{yz}(0, i)],$$

to be the “preference” (possibly negative) of $s_y + s_{yz}$ for setting $\phi(y) = 1$ versus $\phi(y) = 0$, when z has been assigned $\phi(z) = i$. If $\text{bias}_y(i) \leq -1$ then $\phi(y) = 0$ is an optimal assignment. (That is, for every assignment to the remaining variables, including the possibility that $\phi(x) = 0$, setting $\phi(y) = 0$ yields at least as large a score as $\phi(y) = 1$.) Also, if $\text{bias}_y(i) \geq 0$ then $\phi(y) = 1$ is an optimal assignment.

Thus, an optimal assignment $\phi(y)$ can be determined as a function of $\phi(z)$ alone, with no dependence on $\phi(x)$. (This cannot be done in the general case where xy and yz are both heavy edges.) With $\phi(y)$ a function of $\phi(z)$, the score $s_{yz}(\phi(y), \phi(z))$ may be incorporated into the monadic score function $s_z(\phi(z))$. Also, there are only 4 functions from $\{0, 1\}$ to $\{0, 1\}$: as a function of $\phi(z)$, $\phi(y)$ must be the constant function 0 or 1 (in which cases $x \vee y$ can be replaced respectively by a monadic or niladic clause) or $\phi(z)$ or $\overline{\phi(z)}$ (in which cases $x \vee y$ can be replaced respectively by the Sat clause $x \vee z$ or $x \vee \bar{z}$).

This shows that if there is no edge xz and either xy or yz is light, then the 2-reduction produces a light edge xz . If both xy and yz are light, $\nu(F') - \nu(F) \leq -a_2 \leq 0$, while (once again) if one of xy and yz is heavy, $\nu(F') - \nu(F) \leq -w_h + w_e$.

To summarize, no new constraint is imposed by 2-reductions. Also, if either of xy or yz is heavy, then we have not merely that $\nu(F') - \nu(F) \leq 0$ but that $\nu(F') - \nu(F) \leq -w_h + w_e$, and we will take advantage of this later on. \square

5.10. 2-cut. Let $\{x, y\}$ be a 2-cut isolating a set of vertices A , $2 \leq |A| \leq 10$. (The 2-cut reduction extends the 2-reduction, thought of as the case $|A| = 1$.) Similarly to the 1-cut above, for each of the four cases $\phi : \{x, y\} \rightarrow 0, 1$ we may determine the optimal assignments of the vertices in A and the corresponding optimal score; adding this score function to the original dyadic score s_{xy} gives an equivalent instance F' on variables $V \setminus A$. There is nothing new in the technicalities, and we omit them.

In general, $\nu' - \nu$ may be equated with the weight change from deleting the original edge xy if any (guaranteed by (25) not to increase the measure), deleting all vertices in A with their incident half-edges (a change of $-\sum_{v \in A} a_{\text{deg}(v)}$), replacing one half-edge from each of x and y into A with a single heavy edge between x and y (not affecting their degrees, and thus a change of $-w_e + w_h$), then doing half-edge reductions to remove any half-edges remaining from other edges in $\{x, y\} \times A$ (guaranteed by reduction 5.6 not to increase the measure). Thus, $-\sum_{v \in A} a_{\text{deg}(v)} - w_e + w_h \leq -2a_3 - w_e + w_h$, where the second inequality uses that $|A| \geq 2$, all vertices have degree ≥ 3 (a 2-reduction is preferred to this 2-cut reduction), and the values a_i are nondecreasing (see (26)). Thus we can assure that $\nu' - \nu \leq 0$ by

$$-2a_3 - w_e + w_h \leq 0,$$

which is already imposed by (17) and (22). \square

6. SOME USEFUL TOOLS

Before getting down to business, we remark that in treating disjunction and conjunction efficiently, as well as decomposable functions (see reduction 5.5 and Remark 5), the only boolean function our algorithm cannot treat efficiently is exclusive-or. The following remark is surely well known.

Remark 6. *The only non-decomposable two-variable boolean functions are conjunction, disjunction, and exclusive-or.*

Proof. A function $s : \{0, 1\}^2 \mapsto \{0, 1\}$ is characterized by a 2×2 table of 0s and 1s. If the table has rank 1 (or 0), we can decompose s into monadic functions writing $s_{xy}(\phi(x), \phi(y)) = s_x(\phi(x)) + s_y(\phi(y))$. A table with zero or four 1s is a constant function, trivially decomposable. A table with

one 1 is the function $\phi(x) \wedge \phi(y)$, up to symmetries of the table and (correspondingly) negations of one or both variables; similarly a table with three 1s is the function $\phi(x) \vee \phi(y)$. In a table with two 1s, either the 1s share a row or column, in which case the function is decomposable, or they lie on a diagonal, which is (up to symmetries and signs) the function $\phi(x) \oplus \phi(y)$. \square

The property of disjunction and conjunction on which we rely (besides having range $\{0, 1\}$) is that they are monotone in each variable. Obviously exclusive-or is not monotone, and it seems that it cannot be accommodated by our methods.

6.1. Super 2-reduction. Suppose that y is of degree 2 and that its optimal color $C \in \{0, 1\}$ is independent of the colorings of its neighbors x and z , i.e.,

$$(28) \quad (\forall D, E) \quad s_y(C) + s_{yx}(C, D) + s_{yz}(C, E) \\ = \max_{C' \in \{0, 1\}} s_y(C') + s_{yx}(C', D) + s_{yz}(C', E).$$

In that case, $s_y(\phi(y))$ can be replaced by $s_y(C)$ and incorporated into the niladic score, $s_{xy}(\phi(x), \phi(y))$ can be replaced by a monadic score $s'_x(\phi(x)) := s_{xy}(\phi(x), C)$ and combined with the existing s_x , and the same holds for s_{yz} , resulting in an instance with y and its incident edges deleted. \square

A super 2-reduction is better than a usual one since y is deleted, not just contracted.

We will commonly *split* on a vertex u , setting $\phi(u) = 0$ and $\phi(u) = 1$ to obtain instances F_0 and F_1 , and solving both.

Lemma 7. *After splitting a simplified instance F on a vertex u incident to a vertex y of degree 3 whose other two incident edges xy and yz are both light, in at least one of the reduced instances F_0 or F_1 , y is subject to a super 2-reduction.*

Proof. In the clauses represented by the light edges xy and yz , let $b \in \{-2, 0, 2\}$ be the number of occurrences of y minus the number of occurrences of \bar{y} . (As in reduction 5.9, we capitalize on the fact that conjunction and disjunction are both elementwise monotone, and that their scores are limited to $\{0, 1\}$.) Following the fixing of u to 0 or 1 and its elimination, let $\text{bias}_y := s_y(1) - s_y(0)$. Given that F was simplified, the edge uy was not decomposable, so by Remark 5 the value of bias_y in F_0 is unequal to its value in F_1 .

First consider the case $b = 0$. If $\text{bias}_y \geq 1$, the advantage from bias_y for setting $\phi(y) = 1$ rather than 0 is at least equal to the potential loss (at most 1) from the one negative occurrence of y in xy and yz , so the assignment $\phi(y) = 1$ is always optimal. Symmetrically, if $\text{bias}_y \leq -1$ we may set $\phi(y) = 0$. The only case where we cannot assign y is when $\text{bias}_y = 0 = -b/2$.

Next consider $b = 2$. (The case $b = -2$ is symmetric.) If $\text{bias}_y \geq 0$ we can fix $\phi(y) = 1$, while if $\text{bias}_y \leq -2$ we can fix $\phi(y) = 0$. The only case where we cannot assign y is when $\text{bias}_y = -1 = -b/2$.

Thus, we may optimally assign y independent of the assignments of x and z unless $\text{bias}_y = -b/2$. Since bias_y has different values in F_0 and F_1 , in at least one case $\text{bias}_y \neq -b/2$ and we may super 2-reduce on y . \square

6.2. Splitting on vertices of degree 5. Kulikov and Kutzkov [KK07] introduced a clever splitting on vertices of degree 5. Although we will not use it until we address instances of degree 5 in Section 10, we present it here since the basic idea is the same one that went into our 2-reductions: that in some circumstances an optimal assignment of a variable is predetermined. In addition to generalizing from degree 3 to degree 5 (from which the generalization to every degree is obvious), [KK07] also applies the idea somewhat differently.

The presentation in [KK07] is specific to 2-Sat. Reading their result, it seems unbelievable that it also applies to Max 2-CSP as long as the vertex being reduced upon has only light edges (even if its neighbors have heavy edges), but in fact the proof carries over unchanged. For completeness and to make the paper self-contained, we present the generalized result.

Lemma 8 (clause learning). *In a Max 2-CSP instance F , let u be a variable of degree 5, with light edges only, and neighbors v_1, \dots, v_5 . Then there exist “preferred” colors C_u for u and C_i for each neighbor v_i such that a valid splitting of F is into three instances: F_1 with $\phi(u) = C_u$; F_2 with $\phi(u) \neq C_u$, $\phi(v_1) = C_1$; and F_3 with $\phi(u) \neq C_u$, $\phi(v_1) \neq C_1$, and $\phi(v_i) = C_i$ ($\forall i \in \{2, 3, 4, 5\}$).*

Proof. For any coloring $\phi : V \rightarrow \{0, 1\}$, let ϕ^0 and ϕ^1 assign colors 0 and 1 respectively to u , but assign the same colors as ϕ to every other vertex. That is, $\phi^i(u) = i$, and $(\forall x \neq u) \phi^i(x) = \phi(x)$.

What we will prove is that for any assignment ϕ in which at least two neighbors do not receive their preferred colors, $s(\phi^{C_u}) \geq s(\phi)$: the assignment in which u receives its preferred color has score at least as large as that in which it receives the other color, and thus we may exclude the latter possibility in our search. (This may exclude some optimal solutions, but it is also sure to retain an optimal solution; thus this trick will not work for counting, but does work for optimization.) That is, if u and one neighbor (specifically, v_1) do not receive their preferred color, then we may assume that every other neighbor receives its preferred color.

It suffices to show the existence of colors C_u and C_i , $i \in 1, \dots, 5$, such that for any ϕ with $\phi(i) \neq C_i$ for two values of $i \in \{1, \dots, 5\}$, we have $s(\phi^{C_u}) \geq s(\phi)$.

Leave the immediate context behind for a moment, and consider any Max 2-CSP instance F in which some variable u has only light edges, and in them appears N_2^+ times positively and N_2^- times negatively. (As in reduction 5.9 and Lemma 7, we are using the fact that conjunction and disjunction are elementwise monotone.) If $\phi(u) = 0$, the total score s^0 from terms involving u satisfies

$$s_u(0) + N_2^- \leq s^0 \leq s_u(0) + N_2^- + N_2^+,$$

and if $\phi(u) = 1$ the corresponding score s^1 satisfies

$$s_u(1) + N_2^+ \leq s^1 \leq s_u(1) + N_2^+ + N_2^-.$$

From the second inequality in the first line and the first inequality in the second line, if $s_u(1) - s_u(0) \geq N_2^-$ then $s^1 \geq s^0$, and for any coloring ϕ , $s(\phi^1) \geq s(\phi^0)$. Symmetrically, if $s_u(0) - s_u(1) \geq N_2^+$ then ϕ^0 always dominates ϕ^1 . Defining the bias

$$b := s_u(1) - s_u(0),$$

we may thus infer an optimal color for u if $b - N_2^- \geq 0$ or $-b - N_2^+ \geq 0$.

If u has degree 5, $(b - N_2^-) + (-b - N_2^+) = -N_2^- - N_2^+ = -5$, and thus one of these two parenthesized quantities must be at least -2.5 , and by integrality at least -2 . Given the symmetry, without loss of generality suppose that $b - N_2^- \geq -2$. The preferred color for u will be $C_u = 1$.

A small table shows that for any conjunctive or disjunctive clause involving u or \bar{u} and some other variable v_i (which without loss of generality we assume appears positively), there exists a color C_i for v_i (according to the case) such that assigning v_i this color increases $b - N_2^-$ by 1 (either by increasing the bias and leaving N_2^- unchanged, or leaving the bias unchanged and decreasing N_2^-).

original clause	set $\phi(v_i) = C_i =$	resulting clause	change in b	change in N_2^-	change in $b - N_2^-$
$(u \vee v_i)$	0	(u)	+1	0	+1
$(u \wedge v_i)$	1	(u)	+1	0	+1
$(\bar{u} \vee v_i)$	1	(1)	0	-1	+1
$(\bar{u} \wedge v_i)$	0	(0)	0	-1	+1

Thus, starting from $b - N_2^- \geq -2$, assigning to any two neighbors of u their color C_i results in an instance in which $b - N_2^- \geq 0$, and thus in which an optimal assignment for u is $\phi(u) = C_u = 1$. This proves the lemma. \square

6.3. A lemma on 1-reductions. A half-edge reduction or 1-reduction is “good” if the target vertex has degree at least 3, because (as the weights will come out) the measure decrease due to $a_{d-1} - a_d$ is substantial for $d \geq 3$, but small (in fact, 0) for $d = 1$ and $d = 2$.

If for example we start with a simplified instance (in which all vertices must have degree at least 3) and reduce on a vertex of degree d , deleting it and its incident half-edges, each of the d remaining half-edges implies a good degree reduction on a neighboring vertex. However, if we deleted several vertices, this might not be the case: if two deleted vertices had a common neighbor of degree 3, its degree would be reduced from 3 to 2 by one half-edge reduction (good), but then from 2 to 1 by the other (not good).

The following lemma allows us to argue that a certain number of good half-edge reductions occur. The lemma played a helpful role in our thinking about the case analysis, but in the presentation here we invoke it rarely: the cases dealt with are relatively simple, and explicit arguments are about as easy as applying the lemma.

Note that for any half-edge incident on a vertex v , we can substitute a full edge between v and a newly introduced vertex v' : after performing a half-edge reduction on v in the first case or a 1-reduction in the second, the same instance results. (Also, the measure increase of a_1 when we add the degree-1 vertex and half-edge is canceled by the extra decrease for performing a 1-reduction rather than a half-edge reduction.) For clarity of expression, the lemma is thus stated in terms of graphs and 1-reductions, avoiding the awkward half-edges.

Lemma 9. *Let G be a graph with k degree-1 vertices, $X = \{x_1, \dots, x_k\}$. It is possible to perform a series of 1-reductions in G where each vertex x_i in X is either matched one-to-one with a good 1-reduction (a 1-reduction on a vertex of degree 3 or more), or belongs to a component of G containing at least one other vertex of X , where the total order of all such components is at most $2k$ plus the number of degree-2 vertices.*

In particular, if G is a connected graph then there are k good 1-reductions. By analogy with the well-definedness of the 2-core of a graph, any series of 1-reductions should be equivalent, but the weaker statement in the lemma suffices for our purposes.

Proof. The intuition is that each series of reductions originating at some $x_i \in X$, after propagating through a series of vertices of degree 2, terminates either at a vertex of degree 3 or more (reducing its degree), establishing a matching between x and a good reduction, or at another vertex $x_j \in X$, in which case the path from x_i to x_j (or some more complex structure) is a component.

Starting with $i = 1$, let us 1-reduce from x_i as long as possible before moving on to x_{i+1} . That is, if we 1-reduce into a vertex of degree 2 we perform a new 1-reduction from that vertex, terminating when we reach a vertex of degree 1 or degree 3 or more. Rather than deleting an edge with a 1-reduction, imagine that the edges are originally black, and each reduced edge is replaced by a red one (which of course is not available for further 1-reductions).

We assert that just before we start processing any x_i , the red-edged graph has components consisting of vertices *all* of whose edges are red (in which case this is also a component in G itself), and components where all vertices but one *component owner* are all-red, and the component owner has at least 1 red edge and at least 2 black edges. We prove this by induction on i , with $i = 1$ being trivial.

Given that it is true before x_i , we claim that: (1) as we reduce starting with x_i , the reduction sequence is uniquely determined; (2) in the red-edged component including x_i , all vertices are all-red except for a single *active* one; and (3) the sequence on x_i ends when we reduce a vertex that

had at least 3 black edges (matching x_i with this good reduction), or a vertex $x_j \in X$, $j > i$ (in which case we will show that the red component including x_i and x_j is also a component of G itself).

We prove these claims by induction on the step number, the base case again being trivial (x_i itself is active). If we reduce into a vertex v with two black edges (we will say it has *black degree 2*), the next reduction takes us out its other black edge, leaving both red. If v was of degree 2 it is added to x_i 's red component; if not, it must have been a component owner (these are the only mixed-color vertices), and we unite the vertex and its component with x_i 's component. If we reduce into a vertex v with at least 3 black edges, we match x_i with the good reduction on v , and v_i owns x_i 's red component. The only remaining possibility is that we reduce into a vertex with 1 black edge, which can only be a degree-1 vertex x_j (with $j > i$), as there are no mixed-color vertices with 1 black edge. In this case we add x_j to x_i 's component, and terminate the sequence of reductions for x_i without a good reduction. However the red component on x_i now has no black edges on any of its vertices, and is thus a component in the original black graph G .

Starting with the k vertices x_i as initial red components, as we generate the component for x_i , the union of all components is expanded as we pass through (and use up) a (true) degree-2 vertex, left unchanged if we pass through a vertex of higher degree with black degree 2, expanded as we enter a terminal all-black degree-3 vertex, and left unchanged if we terminate at another vertex x_j . Then, recalling that k is the number of degree-1 vertices in X and letting k_2 be the number of degree-2 vertices, the total number of vertices in the union of all components is at most the number of seeds (k), plus the number of pass-throughs (at most k_2), plus the number of good terminals (at most k). In particular, we can partition X into the set of vertices with good terminals in G , and the rest; the rest lie in components of G where the total size of these components is $\leq 2k + k_2$. \square

7. SPLITTING REDUCTIONS AND PREFERENCE ORDER

Recall from Algorithm 1 that if we have a nonempty simplified instance F , we will apply a splitting reduction to produce smaller instances F_1, \dots, F_k , simplify each of them, and argue that $\sum_{i=1}^k 2^{\mu(F_i) - \mu(F)} \leq 1$ (inequality (8')).

We apply splitting reductions in a prescribed order of preference, starting with division into components.

7.1. Split large components. If the constraint graph G of F has components G_1 and G_2 with at least C vertices each (C is the same constant as in the simplification rule (5.4)), decompose F into the corresponding instances F_1 and F_2 . The decomposition is the obvious one: monadic score functions s_x of F are apportioned to F_1 or F_2 according to whether x is a vertex of G_1 or G_2 , similarly for dyadic score functions and edges xy , while we may apportion the niladic score function of F to F_1 , setting that of F_2 to 0.

It is clear that this is a valid reduction, but we must show that (8') is satisfied. Note that $\nu(F_1) + \nu(F_2) = \nu(F)$, and $\nu(F_i) \geq Ca_3$ since F_i has at least C vertices, all degrees are at least 3, and the a_i are nondecreasing. Thus $\nu(F_1) \leq \nu(F) - Ca_3$. Also, $\delta(F_1) - \delta(F)$ is constant-bounded. Assuming that $a_3 > 0$, then for C sufficiently large,

$$\begin{aligned} \mu(F_1) - \mu(F) &= \nu(F_1) - \nu(F) + \delta(F_1) - \delta(F) \\ &\leq -Ca_3 + \sum_{d=4}^6 (|R_d| + |C_d|) \\ &\leq -1. \end{aligned}$$

The same is of course true for F_2 , giving $2^{\mu(F_1) - \mu(F)} + 2^{\mu(F_2) - \mu(F)} \leq 2^{-1} + 2^{-1} = 1$ as required.

The non-strict inequality $a_3 \geq 0$ is established by (17), and if $a_3 = 0$, a 3-regular (cubic) instance would have measure 0, implying that we could solve it in polynomial time, which we do not know how to do. Thus let us assume for a moment that

$$(29) \quad a_3 > 0.$$

This strict inequality (in fact $a_3 \geq 1/7$) will be implied by the constraints for splitting rules for cubic instances, constraint (31) for example. \square

If F 's constraint graph is connected the splitting we apply depends on the degree of F , that is, the degree of its highest-degree vertex. Although high-degree cases thus take precedence, it is easier to discuss the low-degree cases first. Sections 8, 9, 10, and 11 detail the splittings for (respectively) instances of degree 3, 4, 5, and 6. For a given degree, we present the reductions in order of priority.

8. CUBIC INSTANCES

Many formulas are not subject to any of the simplification rules above nor to large-component splitting. In this section we introduce further reductions so that for any formula F of maximum degree at most 3 (which is to say, whose constraint graph has maximum degree at most 3), some reduction can be applied.

If F has any vertex of degree strictly less than 3, we may apply the 0-, 1-, or 2-reductions above. Henceforth, then, we assume that F is 3-regular (cubic).

The new reductions will generally be “atomic” in the sense that we will carry each through to its stated completion, not checking at any intermediate stage whether an earlier simplification or reduction rule can be applied.

We define

$$(30) \quad h_3 := a_3 - a_2$$

to be the decrease of measure resulting from a half-edge reduction (reduction 5.6) on a vertex of degree 3.

8.1. 3-cut. There is a 3-cut $X = \{x_1, x_2, x_3\}$ isolating a set S of vertices, with $4 \leq |S| \leq 10$. Each cut vertex x_i has at least 1 neighbor in $V \setminus \{S \cup X\}$ (otherwise X without this vertex is a smaller cut), and without loss of generality we may assume that either each cut vertex has 2 neighbors in $V \setminus \{S \cup X\}$, or that $|S| = 10$. (If a cut vertex, say x_1 , has just one neighbor $x'_1 \in V \setminus \{S \cup X\}$, then $\{x'_1, x_2, x_3\}$ is also a 3-cut, isolating the larger set $S \cup \{x_1\}$. Repeat until $|S| = 10$ or each cut vertex has two neighbors in $V \setminus \{S \cup X\}$.)

With reference to Figure 2, let $y_1, y_2, y_3 \in S$ be the respective neighbors of x_1, x_2 , and x_3 , and let v_1 and v_2 be the other neighbors of x_1 . Note that $y_2 \neq y_3$, or we should instead apply a 2-cut reduction (reduction 5.10): cutting on $\{x_1, y_2\}$ isolates the set $S \setminus \{y_2\}$, and $3 \leq |S \setminus \{y_2\}| \leq 9$ satisfies the conditions of the 2-cut reduction.

We treat this case by splitting on x_1 , resulting in new instances F_1 and F_2 . In each we apply a 2-cut on $\{y_2, y_3\}$ (not $\{x_2, x_3\}$!), creating a possibly-heavy edge y_2y_3 . We then 2-reduce on y_2 and y_3 in turn to create an edge x_2x_3 which is heavy only if x_2y_2 and x_3y_3 were both heavy. If $|S| \leq 10$, the resulting instances satisfy

$$\mu(F_1), \mu(F_2) \leq \mu(F) - 5a_3 - 2h_3.$$

(Recall that for graphs of degree 3, μ and ν are identical.) The term $-5a_3$ accounts for the deletion of x_1 and S (at least 5 vertices) with their incident half-edges. The term $-2h_3$ accounts for deletion of the “other halves” of the edges from x_1 to $V \setminus \{S \cup X\}$ and the degree decrease of their incident vertices (see definition (30)); we are using the fact that $v_1 \neq v_2$, and that X is an independent set. There is no need for a term accounting for the deletion of the “other halves” of the edges on x_2 and x_3 and the addition of the new edge x_2x_3 : the new x_2x_3 is heavy only if both half-edges were

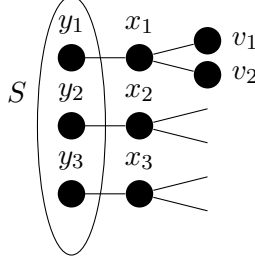


FIGURE 2. Illustration of a 3-cut, reduction 8.1.

heavy, so this change in measure is $-\frac{1}{2}w(x_2y_2) - \frac{1}{2}w(x_3y_3) + w(x_2x_3) \leq 0$, and we are free to ignore it. (Since it may in fact be 0, there is also no gain to including it.) Constraint (8') of Lemma (1) is thus assured if

$$2^{-5a_3-2h_3} + 2^{-5a_3-2h_3} \leq 2^0 = 1.$$

We will henceforth express such constraints by a shorthand, simply saying that the case has splitting number at most

$$(31) \quad (5a_3 + 2h_3, 5a_3 + 2h_3).$$

We formally define a *splitting number* to be

$$(\alpha_1, \alpha_2, \dots, \alpha_k) := 2^{-\alpha_1} + 2^{-\alpha_2} + \dots + 2^{-\alpha_k}.$$

Note the change of sign: in this notation we show the cost *decrease* in each case.

By similar reasoning, if $|S| = 10$ the splitting number is at most

$$(11a_3 + h_3, 11a_3 + h_3).$$

By (29) this constraint is bound to hold “for a sufficiently large value of 10” (and since $h_3 \leq a_3$, for 10 itself this constraint is dominated by (31)), so we will disregard it. \square

8.2. Vertex with independent neighbors. There is a vertex u such that $N(u)$ is an independent set.

With reference to Figure 3, we reduce on u , fixing $\phi(u)$ to 0 and 1 to generate new instances F_0

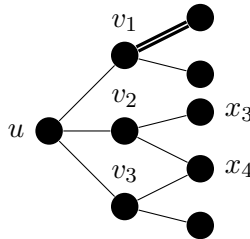


FIGURE 3. Illustration for reduction 8.2, on a vertex with independent neighbors.

and F_1 , each with constraint graph $G[V \setminus \{u\}]$.

Let $N^1 = N(u)$ and $N^2 = N^2(u)$. Let q be the number of vertices in N^1 with a heavy edge to N^2 , k_0 the number of vertices in N^1 subject to a super 2-reduction (deletion) in F_0 , and k_1 the number subject to super 2-reduction in F_1 . By Lemma 7, each $v \in N^1$ falls into at least one of these cases, so $q + k_0 + k_1 \geq 3$.

We will argue that $\mu(F) - \mu(F_i) \geq a_3 + 3h_3 + q(w_h - w_e) + 2k_i h_3$. Deletion of u and reduction of the degree of each of its neighbors immediately reduces the measure by $a_3 + 3h_3$ (more if any

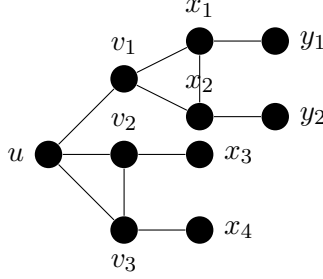


FIGURE 4. Illustration of reduction on a vertex with one edge in its neighborhood, Case 8.3.

edges incident to u were heavy). In F_i , first 2-reduce on the q vertices in N^1 with heavy edges (reducing the measure by a further $q(w_h - w_e)$) and on the $3 - q - k_i$ vertices subject to only plain 2-reductions (not increasing the measure). Note that each vertex in N^2 still has degree 3.

Finally, reduce out the k_i vertices which are set constant by a super 2-reduction, by deleting their incident edges one by one. No vertex v in N^2 has 3 neighbors in N^1 : if it did there would remain only 3 other edges from N^1 to N^2 , whence $|N^2| \leq 4$, $N^2 \setminus v$ would be a cut of size ≤ 3 isolating $N^1 \cup \{u, v\}$, and we would have applied a cut reduction. Thus, deletion of each of the $2k_i$ edges in $N^1 \times N^2$ either reduces the degree of a vertex in N^2 from 3 to 2 (a good 1-reduction, reducing the measure by h_3), or creates a vertex of degree 1.

We wish to show that each degree-1 vertex in the graph $G' = G[V \setminus (\{u\} \cup N^1)]$ must also result in a good 1-reduction, giving the $2k_i h_3$ claimed. Note that $|N^2|$ must be 4, 5, or 6 (if it were smaller we would have applied a cut reduction instead). If $|N^2| = 6$ then every vertex in N^2 has degree 2 (in the graph G') and there is nothing to prove. If $|N^2| = 5$ then at most one vertex in N^2 has degree 1, and Lemma 9 implies that it results in a good 1-reduction. If $|N^2| = 4$, every degree-1 vertex in N^2 also results in a good 1-reduction: If not, then by Lemma 9 a set X of two or more vertices in N^2 lies in a small component of G' , in which case $N^2 \setminus X$ is a cut of size 2 or less in the original constraint graph G , isolating $\{u\} \cup N^1 \cup X$, and we would have applied a cut reduction instead.

Thus, $\mu(F) - \mu(F_i) \geq a_3 + 3h_3 + q(w_h - w_e) + 2k_i h_3$. By convexity, if two splitting numbers have equal total, the more unbalanced one is the more constraining; in this case that means the worst cases come if $k_0 = 0$ and $k_1 = 3 - q$ (or vice-versa). Thus, the worst-case splitting numbers are

$$(32) \quad (\forall q \in \{0, 1, 2, 3\}) \quad (a_3 + 3h_3 + q(w_h - w_e), a_3 + 3h_3 + q(w_h - w_e) + 2(3 - q)h_3).$$

□

8.3. One edge in $G[N(u)]$. Given that we are in this case rather than Case 8.2, no vertex of $N(u)$ has an independent set as neighborhood. Let $N(u) = \{v_1, v_2, v_3\}$ and suppose without loss of generality that $v_2 v_3 \in E$. Let $N(v_1) = \{u, x_1, x_2\}$. Then, $x_1 x_2 \in E$. To avoid a 3-cut (Case 8.1), $|N^2(\{u, v_1\})| = 4$ (the 4 rightmost vertices depicted in Figure 4 are truly distinct).

After splitting on u , in each of the two instances F_0 and F_1 , first 2-reduce on v_1 , then on x_1 , then continue with 2-reductions (the base case), or super 2-reductions (if possible), on v_2 and v_3 . In the base case this results in the deletion of all 5 of these vertices with their incident edges and the decrease of the degree of x_2 to 2, for a measure decrease of $5a_3 + h_3$ (vertex x_2 will be 2-reduced, which does not increase the measure; see 5.9).

If $v_2 v_3$ or $v_2 x_3$ is heavy, then there is an extra measure decrease of $w_h - w_e$ beyond that of the base case, for a splitting number of at most

$$(33) \quad (5a_3 + h_3 + w_h - w_e, 5a_3 + h_3 + w_h - w_e).$$

Otherwise, v_2v_3 and v_2x_3 are both light, and we may super 2-reduce on v_2 in either F_0 or F_1 (without loss of generality say F_1). This reduces the degree of x_3 from 3 to 2, and that of v_3 from 2 to 1, setting up a 1-reduction on v_3 that reduces the degree of x_4 from 3 to 2. This gives a splitting number of at most

$$(34) \quad (5a_3 + h_3, 5a_3 + 3h_3).$$

□

There are no further cases for cubic graphs. If for a vertex u there are 3 edges in $G[N(u)]$ then $N[u]$ is an isolated component (a complete graph K_4) and we apply component-splitting. If there are 2 edges in $G[N(u)]$, then some $v \in N(u)$ (either of the vertices having a neighbor outside $\{u\} \cup N(u)$) has just 1 edge in $G[N(v)]$ and we are back to Case 8.3.

8.4. Cubic results. For results on cubic and other instances, we refer to Theorem 4, Table 2, and the discussion in Section 12.

8.5. Remark on heavy edges. If the original cubic instance is a pure 2-Sat formula, with no heavy edges, then (as we show momentarily) any heavy edges introduced by the procedure we have described can immediately be removed. Thus the “hybrid formula” concept gives no gain for cubic 2-Sat formulas, but expands the scope to cubic Max 2-CSP, sacrifices nothing, and is useful for analyzing non-cubic instances. We now show how heavy edges introduced into a cubic 2-Sat formula immediately disappear again.

In a graph with only light edges, the only two rules that create heavy edges are 2-reductions and 2-cuts (and other reductions that apply these). A 2-reduction on v introduces a heavy edge only if v 's neighbors x_1 and x_2 were already joined by an edge. In that case, though, x_1 and x_2 have their degrees reduced to 2 (at most). If the remaining neighbors y_1 of x_1 and y_2 of x_2 are distinct, then 2-reducing on x_1 gives a light edge x_2y_1 : the heavy edge x_1x_2 is gone. Otherwise, $y_1 = y_2$, and 2-reduction on x_1 followed by 1-reduction on x_2 deletes x_1 and x_2 and reduces the degree of y_2 to 1, again leaving no heavy edge.

For a 2-cut on x_1 and x_2 isolating a set S , if there was an edge x_1x_2 then the cut reduction reduces the degrees of both x_1 and x_2 to 2, and, just as above, we may 2-reduce on x_1 to eliminate the heavy edge. If x_1 and x_2 are nonadjacent and x_1 has just 1 neighbor outside S , then again a follow-up 2-reduction on x_1 eliminates the heavy edge x_1x_2 . Dismissing the symmetric case for x_2 , all that remains is the case when x_1 and x_2 are nonadjacent and each has 2 neighbors outside S , and thus just 1 neighbor in S ; see Figure 5.

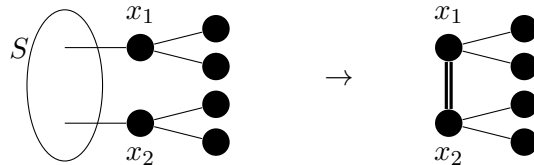


FIGURE 5. 2-cut rule creates a heavy edge.

The S -neighbors x'_1 of x_1 and x'_2 of x_2 must be distinct, or else we would have applied a 1-cut reduction on x'_1 . (This presumes that $|S \setminus \{x'_1\}| \geq 2$, but if it is 0 or 1, we would have 2-reduced on x'_1 or 1-reduced on its S -neighbor — either of which is really a special case of a 1-cut reduction.)

Given that $x'_1 \neq x'_2$, apply a 2-cut reduction not on x_1 and x_2 but instead on x'_1 and x'_2 . Following this with 2-reduction on x'_1 and x'_2 eliminates the heavy edge $x'_1x'_2$, giving a light edge x_1x_2 instead; see Figure 6.

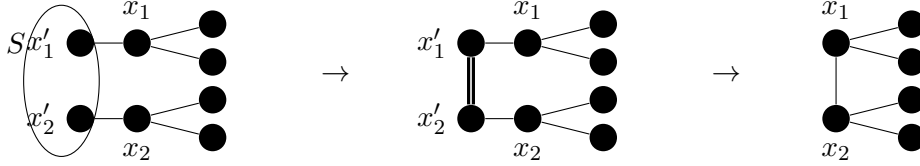


FIGURE 6. 2-cut rule avoids creating a heavy edge.

8.6. Solving the programs. Every weight constraint we introduce is of the form $\sum_i 2^{L_i} \leq 1$, where the sum is finite and each L_i is some linear combination of weights. (Some constraints are simply of the form $L \leq 0$, but this can also be written as $2^L \leq 1$.) This standard form (along with the objective of minimizing w_e) can be provided, through an interface such as AMPL, to a variety of mathematical-programming solvers: we used both IPOPT (part of the free, open-source code repository at `coin-or.org`) and MINOS (a commercial solver).

Furthermore, it is easily verified that the feasible region is convex. (Convexity of 2^x means that for any $p, q \geq 0$, with $p + q = 1$, term by term, $2^{pL+qL'} \leq p2^L + q2^{L'}$, and thus a mixture of feasible solutions is feasible.) This in turn makes it relatively easy for a solver to return a provably optimal solution: convex programs are much easier to solve than general ones or even the quasi-convex programs like Eppstein’s [Epp06].

IPOPT solves the nonlinear program for our general algorithm, to optimality, in a second or two on a typical laptop computer.

To insure that our solutions are truly feasible, in the presence of finite numerical accuracy, we replace the “1” in the right-hand side of each constraint with $1 - \epsilon$, fixing $\epsilon = 10^{-6}$; this allows some margin for error. The values we show for the key parameters w_e and w_h are rounded up (pessimistically) from the higher-precision values returned by the solver, with the other parameter values rounded fairly. Ideally we would also verify, in an unlimited-accuracy tool such as Mathematica, that our rounded values satisfy the original “ ≤ 1 ” constraints, but we have not performed that final check.

9. INSTANCES OF DEGREE 4

We first introduce one more bit of notation, generalizing our earlier definition of h_3 (30). For any $d \geq 3$, we define

$$(35) \quad h_d := \min_{3 \leq i \leq d} \{a_i - a_{i-1}\}.$$

This is the minimum possible decrease of measure resulting from a half-edge reduction (reduction 5.6) on a vertex of degree i with $3 \leq i \leq d$. We will find that such deletions always occur with the same sign in our nonlinear program — the larger h_d , the weaker each constraint is — and therefore the above definition can be expressed in our mathematical program by simple inequalities

$$(36) \quad (\forall 3 \leq i \leq d) \quad h_d \leq a_i - a_{i-1}.$$

We now consider a formula F of (maximum) degree 4. The algorithm chooses a vertex u of degree 4 with — if possible — at least one neighbor of degree 3. The algorithm sets u to 0 and 1, simplifies each instance as much as possible (see Section 5), and recursively solves the resulting instances F_0 and F_1 .

The instances F_0 and F_1 are either 4-regular, of degree at most 3, or nonregular. By the arguments presented in Section 3.6, the case where the degree of the graph decreases can be safely ignored (the measure decrease $C_4 - C_3$ can be made as large as necessary).

9.1. **4-regular.** If F is 4-regular, first consider the case in which F_0 and F_1 are 4-regular. Since splitting on u decreases the degree of each vertex in $N(u)$, and none of our reduction rules increases the degree of a vertex, every vertex in $N(u)$ must have been removed from F_0 and F_1 by simplification rules.² This gives a splitting number of at most

$$(37) \quad (5a_4, 5a_4).$$

If neither F_0 nor F_1 is 4-regular, then u is removed (a_4), the degree of its neighbors decreases ($4h_4$), and we obtain an additional gain because F_0 and F_1 are not regular (R_4). Thus, the splitting number is at most

$$(38) \quad (a_4 + 4h_4 + R_4, a_4 + 4h_4 + R_4).$$

If exactly one of F_0 and F_1 is 4-regular, we obtain a splitting number of $(5a_4, a_4 + 4h_4 + R_4)$. This constraint is weaker (no stronger) than (37) if $5a_4 \leq a_4 + 4h_4 + R_4$, and weaker than (38) if $5a_4 > a_4 + 4h_4 + R_4$, so we may dispense with it.

9.2. **4-nonregular.** If F is not 4-regular, we may assume that u has at least one neighbor of degree 3. Let us denote by p_i the number of degree- i neighbors of u . Thus, $1 \leq p_3 \leq 4$, and $p_3 + p_4 = 4$. Further, let us partition the set P_3 of degree-3 neighbors into those incident only to light edges, P'_3 , and those incident to at least one heavy edge, P''_3 . Define $p'_3 = |P'_3|$ and $p''_3 = |P''_3|$ (so $p'_3 + p''_3 = p_3$).

For each F_i (F_0 and F_1), splitting on u removes u (for a measure decrease of a_4 , compared with F). If F_i is not 4-regular, the degrees of the neighbors of u all decrease ($\sum_{i=3}^4 p_i h_i$). If F_i is regular ($-R_4$), all neighbors of u must have been eliminated as well ($\sum_{i=3}^4 p_i a_i$).

We now argue about additional gains based on the values of p'_3 and p''_3 , starting with the heavy edges incident on vertices in P''_3 . Identify one heavy edge on each such vertex. If such an edge is between two vertices in P''_3 associate it with either one of them; otherwise associate it with its unique endpoint in P''_3 . This gives a set of at least $\lceil p''_3/2 \rceil$ vertices in P''_3 each with a distinct associated heavy edge, which we may think of as oriented out of that vertex. If such an edge incident on $v \in P''_3$ is also incident on u then it is deleted along with u , for an additional measure reduction of $w_h - w_e$ we credit to v . This leaves a set of “out” edges that may form paths or cycles. After deletion of u all the vertices involved have degree 2, so any cycle is deleted as an isolated component, for a measure reduction of $w_h - w_e$ per vertex. Super 2-reducing on a vertex v deletes its outgoing edge, which we credit to v , and possibly also an incoming heavy edge associated with a different $v' \in P''_3$, which we credit to v' . Finally, if v is 2-reduced we consider its outgoing edge (not its other incident edge) to be contracted out along with v , crediting this to v (and correctly resulting in a light edge if the other edge incident on v was light, or a heavy one if it was heavy). This means that if the other edge incident to v was a heavy edge out of a different $v' \in P''_3$, then v' still has an associated outgoing heavy edge. In short, each of the $\lceil p''_3/2 \rceil$ vertices gets credited with the loss of a heavy edge, for an additional measure reduction of at least $\lceil p''_3/2 \rceil (w_h - w_e)$.

We say that we have a *good degree reduction* if the degree of a vertex of degree 3 or more decreases by 1: for graphs of degree 4 this decreases the measure by at least h_4 . This measure

²There is an important subtlety here: the reduced-degree vertices are eliminated, not merely split off into other components such that F_i has a 4-regular component and a component of degree 3 (although such an example shares with 4-regularity the salient property that no degree-4 vertex has a degree-3 neighbor). By definition, the “4-regular case” we are considering at this point does not include such an F_i , but it is worth thinking about what happens to an F_i which is not regular but has regular components. No component of F_i is small (simplification 5.4 has been applied), so in the recursive solution of F_i , Algorithm 1 immediately applies large-component splitting (reduction 7.1). This reduces F_i to two connected instances, and is guaranteed to satisfy constraint (8') (the penalty for one instance's being 4-regular is more than offset by its being much smaller than F_i). Our machinery takes care of all of this automatically, but the example illustrates why some of the machinery is needed.

decrease comes in addition to what we have accounted for so far, unless F_i is regular and the degree reduction is on a vertex in $N(u)$ (since we have accounted for the deletion of those vertices, counting their degree reductions as well would be double counting). We will show that a certain number of additional-scoring degree reductions occur altogether, in F_0 and F_1 combined, as a function of p'_3 .

If $p'_3 = 1$, super 2-reduction on the sole vertex in P'_3 is possible in at least one of F_0 or F_1 — without loss of generality say just F_0 — and reduces the degrees of at least two neighbors. If F_0 is nonregular this gives a gain of $2h_4$, while if F_0 is regular there may be no gain.

If $p'_3 = 2$, then again if either vertex is super 2-reduced in a nonregular branch there is a gain of at least $2h_4$. Otherwise, each vertex is super 2-reduced in a regular branch (both in one branch, or in two different branches, as the case may be). At least one of the vertices has at least one neighbor in $N^2 := N^2(G)$, or else $P_3 \setminus P'_3$ would be 2-cut. In whichever F_i the degree of the neighbor is reduced, since F_i is regular the neighbor must eventually be deleted, for a gain of at least a_3 . So there is either a gain of $2h_4$ in a nonregular branch or a gain of a_3 in a regular branch. (We cannot hope to replace a_3 with $2a_3$: Figure 7 shows an example where indeed only one good degree reduction occurs outside $N[u]$.)

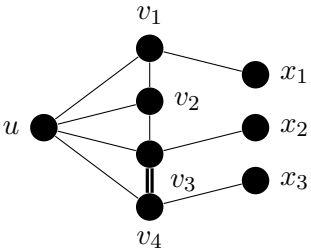


FIGURE 7. The case $p'_3 = 2$ may lead to just one good degree reduction outside $N[u]$. If both super 2-reductions on v_1 and v_2 occur in the same branch (say F_1), the degree of x_1 is reduced. The degrees of v_3 and v_4 become 2, so their edges are contracted eventually creating an edge x_2x_3 , which does not change the degree of x_2 or x_3 . The heavy edge v_3v_4 gives a bonus measure reduction of $w_h - w_e$ previously accounted for.

If $p'_3 = 3$, again either there is a gain of $2h_4$ in a nonregular branch, or each super 2-reduction occurs in a regular branch. The 3 vertices in P'_3 have at least 2 neighbors in N^2 , or else these neighbors, along with $P_3 \setminus P'_3$, would form a cut of size 2 or smaller. Each of these neighbors has its degree reduced, and thus must get deleted from a regular F_i , for a gain of at least $2a_3$. So there is either a gain of $2h_4$ in a nonregular branch, or a gain of $2a_3$ altogether in one or two regular branches. (We cannot hope to claim $3h_4$ or $3a_3$, per the example in Figure 8.)

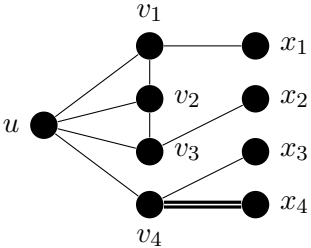


FIGURE 8. The case $p'_3 = 3$ ($P'_3 = \{v_1, v_2, v_3\}$) may lead to just two good degree reductions.

If $p'_3 = 4$, we claim that in the two branches together there are at least 4 good degree reductions on vertices in N^2 and $N^3(u)$. Each contributes a gain of at least h_4 if it is in a nonregular branch, a_3 in a regular branch. Each vertex in N^2 undergoes a good degree reduction in one branch or the other, so if $|N^2| \geq 4$ we are done. Since there can be no 2-cut, we may otherwise assume that $|N^2| = 3$. Since (in F) every vertex in $N(u)$ has degree 3, there is an even number of edges between $N(u)$ and N^2 , thus there are at least 4 such edges. Since each vertex in N^2 has an edge from $N(u)$, there must be two such edges incident on one vertex $x_1 \in N^2$, and one edge each incident on the other vertices $x_2, x_3 \in N^2$. Again we guaranteed 4 good degree reductions unless x_1 has degree 3 *and* undergoes both of its reductions in one branch (so that degree 3 to 2 is a good reduction, but 2 to 1 is not). In that case, though, x_1 has degree 1, its remaining neighbor must be in $N^3(u)$ (otherwise $\{x_1, x_2\}$ is a 2-cut), and 1-reducing on x_1 gives a good degree reduction on that neighbor. So there is a total gain of $4h_4$ in a nonregular branch and $4a_3$ in a regular branch.

By convexity, the elementwise average of two pairs of splitting numbers is a constraint dominated by one or the other, so it suffices to write down the extreme constraints, with all the gain from super 2-reductions given to a single nonregular or regular branch.

Before counting the super 2-reduction gains, if F_i is nonregular the measure decrease $\mu(F) - \mu(F_i)$ is at least

$$(39) \quad \Delta_{\bar{r}}(p_3, p''_3, p_4) := a_4 + \sum_{i=3}^4 p_i h_i + \left\lceil \frac{p''_3}{2} \right\rceil (w_h - w_e),$$

and if F_i is 4-regular, at least

$$(40) \quad \Delta_r(p_3, p''_3, p_4) := a_4 + \sum_{i=3}^4 p_i a_i + \left\lceil \frac{p''_3}{2} \right\rceil (w_h - w_e) - R_4.$$

The super 2-reductions give an additional gain, in a nonregular branch, of at least

$$(41) \quad g_{\bar{r}} := \left\lfloor \frac{p'_3 + 2}{3} \right\rfloor 2h_4,$$

and in a regular branch, at least

$$(42) \quad g_r := \left(\left\lfloor \frac{p'_3}{2} \right\rfloor + \left\lfloor \frac{p'_3}{3} \right\rfloor + \left\lfloor \frac{p'_3}{4} \right\rfloor \right) a_3,$$

where the tricky floor and ceiling expressions are just a way of writing an explicit expression convenient for passing to the nonlinear solver. The constraints arising from splitting on a vertex of degree 4 with at least one neighbor of degree 3 are thus dominated by the following, taken over $p'_3 + p''_3 + p_4 = 4$, with $p_4 \leq 3$ and $p_3 = p'_3 + p''_3$:

$$(43) \quad (\Delta_{\bar{r}}, \Delta_{\bar{r}} + g_{\bar{r}}),$$

$$(44) \quad (\Delta_{\bar{r}}, \Delta_r + g_r),$$

$$(45) \quad (\Delta_r, \Delta_{\bar{r}} + g_{\bar{r}}),$$

$$(46) \quad (\Delta_r, \Delta_r + g_r)$$

10. INSTANCES OF DEGREE 5

This section considers formulas of maximum degree 5. As an overview, if there is a 3-cut isolating a set S with 6 or more vertices the algorithm splits on any vertex in the cut. Otherwise, the algorithm chooses a vertex u of degree 5 with — if possible — at least one neighbor of degree at most 4, and splits on u either as was done in the degree-4 case, or using clause-learning splitting (see Lemma 8). We use clause learning when the neighbors of u have high degrees, because clause learning sets many variables in $N(u)$, and this is most effective when the degrees are large (since

$a_i \geq a_{i-1}$). We use normal splitting when the neighbors have low degrees, because setting u reduces their degrees, and this is effective when the degrees are small ($h_i \leq h_{i+1}$, with an additional bonus in super 2-reductions for a degree-3 variable). (This is also why we always prefer to split on vertices of maximum degree with neighbors of low degree, and why the regular cases need special attention.)

10.1. 3-cut. There is a 3-cut $C = \{x_1, x_2, x_3\}$ isolating a set S of vertices such that $6 \leq |S| \leq 10$ and S contains at least one vertex of degree 5. Splitting on the cut vertex x_1 leaves constraint graphs where $\{x_2, x_3\}$ form a 2-cut. Thus $S \cup \{x_1\}$ are removed from both resulting instances ($a_5 + 6a_3$), a neighbor of x_1 outside $S \cup C$ has its degree reduced (h_5), a heavy edge x_2x_3 appears (in the worst case) but at least 2 half-edges incident on x_2 and x_3 disappear ($-w_h + w_e$). Additionally, the resulting instances may become 5-regular ($-R_5$). So, the splitting number is at most

$$(47) \quad (a_5 + 6a_3 + h_5 - w_h + w_e - R_5, a_5 + 6a_3 + h_5 - w_h + w_e - R_5).$$

□

In light of reduction 10.1 we may henceforth assume that each degree-5 variable u has $|N^2(u)| \geq 4$.

10.2. 5-regular. If every vertex has degree 5, the same analysis as for 4-regular instances (reduction 9.1, constraints (37) and (38)) gives a splitting number which is at most one of the following:

$$(48) \quad (6a_5, 6a_5),$$

$$(49) \quad (a_5 + 5h_5 + R_5, a_5 + 5h_5 + R_5).$$

□

Otherwise, let u be a degree-5 vertex with a minimum number of degree-5 neighbors, and as usual let p_i be the number of degree- i neighbors of u (since the instance is not regular, $p_5 < 5$). Let $H := \chi(u \text{ is incident to a heavy edge})$. Depending on the values of H and p_i we will use either regular 2-way splitting (reduction 10.3) or clause-learning 3-way splitting (reduction 10.4).

10.3. 5-nonregular, 2-way splitting. $H = 1$ or $p_3 \geq 1$ or $p_5 \leq 2$.

In this case we use the usual 2-way splitting, setting u to 0 and to 1, and simplifying to obtain F_0 and F_1 . If F_i is not regular, the measure decrease $\mu(F) - \mu(F_i)$ is at least $a_5 + \sum_{i=3}^5 p_i h_i + H(w_h - w_e)$, and if F_i is 5-regular, it is at least $a_5 + \sum_{i=3}^5 p_i a_i + H(w_h - w_e) - R_5$. Thus if both branches are regular the splitting number is at most

$$(50) \quad (a_5 + \sum_{i=3}^5 p_i a_i + H(w_h - w_e) - R_5, a_5 + \sum_{i=3}^5 p_i a_i + H(w_h - w_e) - R_5),$$

and if one branch is regular and one nonregular, at most

$$(51) \quad (a_5 + \sum_{i=3}^5 p_i a_i + H(w_h - w_e) - R_5, a_5 + \sum_{i=3}^5 p_i h_i + H(w_h - w_e)).$$

If both branches are nonregular, we use that if $p_3 \geq 1$, any degree-3 neighbor of u either has a heavy edge not incident to u , giving an additional measure reduction of at least $w_h - w_e$, or in at least one branch may be super 2-reduced, for a measure reduction of at least $2h_5$. (The latter requires a justification we give explicitly, although Lemma 9 could be invoked. At the start of the first super 2-reduction, every vertex has degree 2 or more. Each of the two “legs” of the super 2-reduction propagates through a [possibly empty] chain of degree-2 vertices before terminating either in a good degree reduction or by meeting a vertex that was reduced to degree 1 by the other leg. In the latter case all the vertices involved had degree 2, thus were neighbors of u originally of degree 3; also, there must have been at least three of them to form a cycle, and the remaining 2 or fewer vertices in $N(u)$ contradict the assumption that F was simplified.) Thus, the splitting number is at most

$$(52) \quad (a_5 + \sum_{i=3}^5 p_i h_i + H(w_h - w_e) + \chi(p_3 \geq 1)2h_5, a_5 + \sum_{i=3}^5 p_i h_i + H(w_h - w_e)) \text{ or}$$

$$(53) \quad \begin{aligned} & (a_5 + \sum_{i=3}^5 p_i h_i + H(w_h - w_e) + \chi(p_3 \geq 1)(w_h - w_e), \\ & a_5 + \sum_{i=3}^5 p_i h_i + H(w_h - w_e) + \chi(p_3 \geq 1)(w_h - w_e)). \end{aligned}$$

□

10.4. **5-nonregular, clause learning.** $H = 0$ and $p_3 = 0$ and $p_5 \in \{3, 4\}$.

Let v be a degree 5 (degree 5 in G) neighbor of u with a minimum number of degree-5 neighbors in $N^2 := N^2(u)$. The clause learning splitting (see Lemma 8) will set u in the first branch, u and v in the second branch, and all of $N[u]$ in the third branch. In each of the 3 branches, the resulting instance could become 5-regular or not.

In the *first branch*, the measure of the instance decreases by at least

$$(54) \quad \Delta_{51} := \min \begin{cases} a_5 + \sum_{i=4}^5 p_i h_i & \text{(5-nonregular case), or} \\ a_5 + \sum_{i=4}^5 p_i a_i - R_5 & \text{(5-regular case).} \end{cases}$$

In the analysis of the second and third branches we distinguish between the case where v has at most one neighbor of degree 5 in N^2 , and the case where v (and thus every degree-5 neighbor of u) has at least two neighbors of degree 5 in N^2 .

In the *second branch*, if v has at most one neighbor of degree 5 in N^2 , the measure of the instance decreases by at least

$$(55) \quad \Delta_{52}^1 := \min \begin{cases} a_5 + \sum_{i=4}^5 p_i h_i + a_4 + 3h_4 + h_5 & \text{(5-nonregular case), or} \\ a_5 + \sum_{i=4}^5 p_i a_i - R_5 & \text{(5-regular case).} \end{cases}$$

(The degree reductions $3h_4 + h_5$ from the nonregular case do not appear in the regular case because they may pertain to the same vertices as the deletions $\sum p_i a_i$.)

If v has at least two neighbors of degree 5 in N^2 , the measure decreases by at least

$$(56) \quad \Delta_{52}^2 := \min \begin{cases} a_5 + \sum_{i=4}^5 p_i h_i + a_4 + 4h_5 & \text{(5-nonregular case), or} \\ a_5 + \sum_{i=4}^5 p_i a_i + 2a_5 - R_5 & \text{(5-regular case).} \end{cases}$$

In the *third branch*, first take the case where v has at most one neighbor of degree 5 in N^2 . Since $|N^2| \geq 4$, there are at least 4 good degree reductions on vertices in N^2 . If the instance becomes regular, this implies a measure decrease of at least $4a_3$. If the instance remains nonregular, this is a measure reduction of at least $4h_5$, and we now show that if $p_5 = 4$ then there is a fifth good degree reduction. We argue this just as the 4-nonregular case (Section 9.2) with $p'_3 = 4$; we could alternatively apply Lemma 9. If $|N^2| = 5$ the desired $5h_5$ is immediate. Otherwise, $|N^2| = 4$, the number of edges between $N(u)$ and N^2 is at least 4, and odd (from $p_5 = 4$ and $p_4 = 1$, recalling that $p_3 = 0$), so $|N(u) \times N^2| \geq 5$. At least one edge incident on each vertex in N^2 gives a good degree reduction, and we fail to get a fifth such reduction only if the fifth edge is incident on a vertex $x \in N^2$ of degree 3, leaving it with degree 1. But in that case the remaining neighbor of x must be in $N^3(u)$ (otherwise $N^2 \setminus x$ is a 3-cut, a contradiction by reduction 10.1), and 1-reducing x gives the fifth good degree reduction. Thus the measure decreases by at least

$$(57) \quad \Delta_{53}^1 := \min \begin{cases} a_5 + \sum_{i=4}^5 p_i a_i + 4h_5 + \chi(p_5 = 4)h_5 & \text{(5-nonregular case), or} \\ a_5 + \sum_{i=4}^5 p_i a_i + 4a_3 - R_5 & \text{(5-regular case).} \end{cases}$$

Otherwise, in the third branch, v has at least two neighbors of degree 5 in N^2 . For the regular case we simply note that each vertex in N^2 has its degree reduced and must be deleted, N^2 has at least four vertices of which at least two are of degree 5, for a measure reduction of at least $2a_5 + 2a_3$. We now address the nonregular case. Letting P_5 be the set of degree-5 vertices in $N(u)$ (so $|P_5| = p_5$), by definition of v every vertex in P_5 has at least two degree-5 neighbors in N^2 . Let $R \subseteq N^2$ be the set of degree-5 vertices in N^2 adjacent to P_5 , and let $E_5 = E \cap (P_5 \times R)$ be the

set of edges between P_5 and R . There is one last case distinction, according to the value of p_5 . If $p_5 = 3$ there are at least 6 good degree reductions: $|E_5| = 6$, each vertex in R has at most $|P_5| = 3$ incident edges from E_5 , and thus each such incidence results in a good degree reduction (the vertex degree is reduced at most from 5 to 4 to 3 to 2). Here we have $6h_5$.

If $p_5 = 4$ we claim that the good degree reductions amount to at least $\min\{8h_5, 5h_5 + h_4 + h_3\}$. By default the 8 edges in E_5 all generate good degree reductions, with fewer only if some of the degree-5 vertices in R have more than 3 incident edges from E_5 . The “degree spectrum” on R is thus a partition of 8 (the number of incident edges) into $|R|$ parts, where no part can be larger than $|P_4| = 4$. If the partition is 4 + 4 this means two reductions that are not good ($2h_2$), but then this implies that $|R| = 2$, and the other two vertices in $N^2 \setminus R$ also have their degrees reduced, restoring the total of 8 good reductions. If the partition has exactly one 4, on a vertex $r \in R$, then just one of the 8 degree reductions is not good, and the 7 good reductions include those on r , thus giving a measure reduction of at least $5h_5 + h_4 + h_3$.

Considering the difference, which we will denote $g_{p_5=4}$, between these guaranteed measure decreases and the guarantee of $6h_5$ when $p_5 = 3$, we constrain

$$(58) \quad g_{p_5=4} \leq 8h_5 - 6h_5 = 2h_5,$$

$$(59) \quad g_{p_5=4} \leq (5h_5 + h_4 + h_3) - 6h_5 = -h_5 + h_4 + h_3.$$

and we obtain a measure reduction of at least

$$(60) \quad \Delta_{53}^2 := \min \begin{cases} a_5 + \sum_{i=4}^5 p_i a_i + 6h_5 + \chi(p_4 = 1)g_{p_5=4} & \text{(5-nonregular case), or} \\ a_5 + \sum_{i=4}^5 p_i a_i + 2a_5 + 2a_3 - R_5 & \text{(5-regular case).} \end{cases}$$

Wrapping up this reduction, the case that v has at most 1 degree-5 neighbor in N , or at least two such neighbors, respectively impose the constraints (splitting numbers)

$$(61) \quad (\Delta_{51}, \Delta_{52}^1, \Delta_{53}^1) \text{ and}$$

$$(62) \quad (\Delta_{51}, \Delta_{52}^2, \Delta_{53}^2).$$

□

11. INSTANCES OF DEGREE 6

This section considers formulas of maximum degree 6. The algorithm chooses a vertex u of degree 6 with — if possible — at least one neighbor of lower degree, and splits on u by setting it to 0 and 1.

11.1. **6-regular.** If every vertex has degree 6, the same analysis as for regular instances of degree 4 gives a splitting number which is at least one of the following:

$$(63) \quad (7a_6, 7a_6),$$

$$(64) \quad (a_6 + 6h_6 + R_6, a_6 + 6h_6 + R_6).$$

□

11.2. **6-nonregular.** Vertex u has at least one neighbor of degree at most 5.

It is straightforward that the splitting number is at least as large as one of the following (only distinguishing if the instance becomes 6-regular or not):

$$(65) \quad \left(a_6 + \sum_{i=3}^6 p_i h_i, a_6 + \sum_{i=3}^6 p_i h_i \right),$$

$$(66) \quad \left(a_6 + \sum_{i=3}^6 p_i a_i - R_6, a_6 + \sum_{i=3}^6 p_i a_i - R_6 \right).$$

□

12. TUNING THE BOUNDS

For any values of w_e and w_h satisfying the constraints we have set down, we have shown that any Max 2-CSP instance F is solved in time $O^*(2^{|E|w_e + |H|w_h})$.

For a given instance F , the running-time bound is best for the feasible values of w_e and w_h which minimize $|E|w_e + |H|w_h$. As usual taking $|E| = (1 - p)m$ and $|H| = pm$, this is equivalent to minimizing

$$(67) \quad (1 - p)w_e + pw_h,$$

allowing us to obtain a 1-parameter family of running-time bounds — pairs (w_e, w_h) as a function of p — tuned to a formula’s fraction of conjunctive and general 2-clauses.

Reiterating, if a formula’s “p” value is $p(F) = |H|/(|E| + |H|)$, and if minimizing (67) for a given p gives a pair $(w_e, w_h)(p)$, then the optimal bound for formula F is the one given by $(w_e, w_h)(p(F))$, but for *any* $(w_e, w_h)(p)$, the running-time bound $O^*(2^{|E|w_e + |H|w_h})$ is valid for every formula F , even if $p \neq p(F)$. This is simply because every such pair (w_e, w_h) is a feasible solution of the nonlinear program, even if it is not the optimal solution for the appropriate objective function.

For cubic instances, minimizing (67) with p small gives $w_e \approx 0.10209$ and $w_h \approx 0.23127$, while minimizing with p close to 1 gives $w_e = w_h = 1/6$ (the tight constraints are all linear, so the solution is rational), matching the best known polynomial space running time for general instances of Max 2-CSP (see [SS07a]). It appears that the first result is obtained for all $p \leq 1/2$ and the second for all $p > 1/2$.

For instances of degrees 4, 5, and 6 or more, the results of minimizing with various values of p are shown in Table 2, and the most interesting of these is surely that of degree 6 or more (the general case). Here, taking p small gives $w_e \approx 0.15820$ and $w_h \approx 0.31174$. For instances of Max 2-Sat this gives a running-time bound of $O^*(2^{0.1582m})$ or $O^*(2^{m/6.321})$, improving on the best bound previously known, giving the same bound for mixtures of OR and AND clauses, and giving nearly as good run times when a small fraction of arbitrary integer-weighted clauses are mixed in. We observe that any $p \geq 0.29$ leads to $w_e = w_h = 0.19$ (as for cubic case with $p > 1/2$, the tight constraints are linear, so the value is rational), matching the best known bound (for polynomial-space algorithms) of $O^*(2^{0.19m})$ from [SS07a]. Figure 1 shows the values of w_e , w_h , and the objective $(1 - p)w_e + (p)w_h$, as a function of p . Numerically, the values w_e and w_h meet for some value of p between 0.2899 and 0.29.

13. CONCLUSIONS

In this paper, we introduced a hybrid model for mixed 2-Sat and 2-CSP instances. It captures the algorithmic benefits of both a more general setting, where powerful reductions can be performed, and a narrower one, where more specialized reductions come into play. Whereas in all previous improvements for Max 2-Sat, the specific 2-Sat setting and the more general 2-CSP setting were

permanently competing, we use in this paper, for the first time, a unified setting taking advantage of both worlds.

We introduced hybrid instances in the context of Max 2-Sat and Max 2-CSP, a pair of well-studied and important problems, but we believe they will be of use in branching algorithms for other problems. Natural candidates are k -coloring and $(k,2)$ -CSP, Maximum Independent Set and Max 2-CSP, or Maximum Cut and Max 2-CSP. A virtue of the hybrid approach is that the method itself imposes no overhead. For Max 2-Sat, for example, if we included no transformation producing heavy edges from light ones, the method would yield the same result as an analysis confined to simple clauses. If there is even one rule producing heavy edges, and one rule decreasing the number of such edges, it is already possible to get a better result. The hybrid method could well be generalized beyond light and heavy edges. For example, better running-time bounds for Max 2-Sat and/or Max Cut might come from allowing one type of edges for conjunctions and disjunctions, a second type for exclusive-or, and a third type for general, weighted, 2-variable constraints. In this case it is not clear which of the first two types of constraints would get a heavier weight, but (given a set of transformation rules) the hybrid method will establish the “exchange rate” between all the edge types.

Recently, Bonsma and Lokshtanov [BL] explored feedback vertex set in what they call “mixed” graphs, with directed and undirected edges, and showed that the problem is fixed-parameter tractable. It is interesting to see the hybrid viewpoint adopted but pursued with different techniques.

ACKNOWLEDGMENT

The authors are very grateful to Alex Scott for initiating this project and contributing some of the first key ideas.

REFERENCES

- [BE05] Richard Beigel and David Eppstein, *3-coloring in time $O(1.3289^n)$* , J. Algorithms **54** (2005), no. 2, 168–204.
- [BL] Paul Bonsma and Daniel Lokshtanov, *Feedback vertex set in mixed graphs*, Proceedings of the 12th Algorithms and Data Structures Symposium (WADS 2011), to appear, preprint available at <http://arxiv.org/abs/1010.5974>.
- [BR99] Nikhil Bansal and Venkatesh Raman, *Upper bounds for MaxSat: Further improved*, Proceedings of the 10th International Symposium on Algorithms and Computation (ISAAC 1999), Lecture Notes in Comput. Sci., vol. 1741, Springer, 1999, pp. 247–258.
- [Epp06] David Eppstein, *Quasiconvex analysis of multivariate recurrence equations for backtracking algorithms*, ACM Trans. Algorithms **2** (2006), no. 4, 492–509.
- [FGK09] Fedor V. Fomin, Fabrizio Grandoni, and Dieter Kratsch, *A measure & conquer approach for the analysis of exact algorithms*, J. ACM **56** (2009), no. 5, 1–32.
- [GHN03] Jens Gramm, Edward A. Hirsch, Rolf Niedermeier, and Peter Rossmanith, *Worst-case upper bounds for MAX-2-SAT with an application to MAX-CUT*, Discrete Appl. Math. **130** (2003), no. 2, 139–155.
- [GS09] Serge Gaspers and Gregory B. Sorkin, *A universally fastest algorithm for Max 2-Sat, Max 2-CSP, and everything in between*, Proceedings of the 20th annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2009), SIAM, 2009, pp. 606–615.
- [Hir00] Edward A. Hirsch, *A new algorithm for MAX-2-SAT*, Proceedings of the 17th Annual Symposium on Theoretical Aspects of Computer Science (STACS 2000), Lecture Notes in Comput. Sci., vol. 1770, Springer, 2000, pp. 65–73.
- [KK06] Arist Kojevnikov and Alexander S. Kulikov, *A new approach to proving upper bounds for MAX-2-SAT*, Proceedings of the 17th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2006), ACM, 2006, pp. 11–17.
- [KK07] Alexander S. Kulikov and Konstantin Kutzkov, *New bounds for MAX-SAT by clause learning*, Proceedings of the 2nd International Symposium on Computer Science in Russia (CSR 2007), Lecture Notes in Comput. Sci., vol. 4649, Springer, 2007, pp. 194–204.

- [KMRR05] Joachim Kneis, Daniel Mölle, Stefan Richter, and Peter Rossmanith, *Algorithms based on the treewidth of sparse graphs*, Proceedings of the 31st International Workshop on Graph-Theoretic Concepts in Computer Science (WG 2005), Lecture Notes in Comput. Sci., vol. 3787, Springer, 2005, pp. 385–396.
- [Koi06] Mikko Koivisto, *Optimal 2-constraint satisfaction via sum-product algorithms*, Information Proc. Lett. **98** (2006), no. 1, 24–28.
- [KR05] Joachim Kneis and Peter Rossmanith, *A new satisfiability algorithm with applications to Max-Cut*, Tech. Report AIB-2005-08, Department of Computer Science, RWTH Aachen, 2005.
- [Kul97] Oliver Kullmann, *Worst-case analysis, 3-SAT decision and lower bounds: Approaches for improved SAT algorithms*, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, vol. 35, American Mathematical Society, 1997, pp. 261–313.
- [Kul99] ———, *New methods for 3-SAT decision and worst-case analysis*, Theoret. Comput. Sci. **223** (1999), no. 1-2, 1–72.
- [NR00] Rolf Niedermeier and Peter Rossmanith, *New upper bounds for maximum satisfiability*, J. Algorithms **36** (2000), no. 1, 63–88.
- [RF08] Daniel Raible and Henning Fernau, *A new upper bound for Max-2-SAT: A graph-theoretic approach*, Proceedings of the 33rd International Symposium on Mathematical Foundations of Computer Science (MFCS 2008), Lecture Notes in Comput. Sci., vol. 5162, Springer, 2008, pp. 551–562.
- [SS03] Alexander D. Scott and Gregory B. Sorkin, *Faster algorithms for MAX CUT and MAX CSP, with polynomial expected time for sparse instances*, Proceedings of the 7th International Workshop on Randomization and Approximation Techniques in Computer Science (RANDOM 2003), Lecture Notes in Comput. Sci., vol. 2764, Springer, 2003, pp. 382–395.
- [SS04] ———, *A faster exponential-time algorithm for Max 2-Sat, Max Cut, and Max k-Cut*, Tech. Report RC23456 (W0412-001), IBM Research Report, December 2004, See <http://domino.research.ibm.com/library/cyberdig.nsf>.
- [SS07a] ———, *Linear-programming design and analysis of fast algorithms for Max 2-CSP*, Discrete Optim. **4** (2007), no. 3-4, 260–287.
- [SS07b] ———, *Polynomial constraint satisfaction: A framework for counting and sampling CSPs and other problems*, Tech. Report cs:DM/0604079v3, arxiv.org, February 2007, See <http://arxiv.org/abs/cs.DM/0604079>.
- [Wah04] Magnus Wahlström, *Exact algorithms for finding minimum transversals in rank-3 hypergraphs*, J. Algorithms **51** (2004), no. 2, 107–121.
- [Wil05] Ryan Williams, *A new algorithm for optimal 2-constraint satisfaction and its implications*, Theoret. Comput. Sci. **348** (2005), no. 2-3, 357–365.

APPENDIX: CONVEX PROGRAM FOR COMPUTING THE OPTIMAL WEIGHTS

Below we show, in AMPL notation, the objective function and all the constraints of the mathematical program we solve to optimize an algorithm for hybrid instances with a fraction p of non-simple clauses. Constraints are annotated the numbers of the corresponding inequalities in the paper's body. The parameter `margin` is the " ϵ " discussed in Section 8.6 to ensure that a solution is truly feasible even in the face of finite-precision arithmetic.

```
# Max 2-Sat and Max 2-CSP

# maximum degree
param maxd integer >=3;
# fraction of non-simple clauses
param p;
param margin;
set DEGREES := 0..maxd;
# weight for edges
var We >= 0;
# weight for degree reductions from degree at most i
var h {DEGREES} >= 0;
# vertex of degree i + i/2 surrounding half-edges
var a {DEGREES};
# weight for heavy edges
var Wh;
# Regular weights
var R4 >= 0;
var R5 >= 0;
var R6 >= 0;
# additional degree reductions in the 3rd branch (nonregular)
# of the clause learning branching for p5=4 vs p5=3
var nonreg53;
# change in measure for the 3 branches
# 1st argument is the nb of deg-4 nbs of u
# 2nd argument distinguishes (if present) if v has at most 1 deg-5 nb in N^2 (1)
# or at least 2 (2)
set TWO := 1..2;
var f1 {TWO};
var f2 {TWO,TWO};
var f3 {TWO,TWO};
var D4r {0..4, 0..4};
var D4n {0..4, 0..4};
var g4r {0..4};
var g4n {0..4};

# analysis in terms of the number of edges
minimize Obj: (1-p)*We + p*Wh;

# Some things we know
subject to Known:
    a[0] = 0;
```

(14)

(14)

(14)

(1)

(2)

(24)

```
# Constrain W values non-positive
subject to Wnonpos {d in DEGREES : d>=1}:
    a[d] - d*We/2 <= 0 - margin; (16)(12)
```

```
# a[] value positive
subject to MeasurePos {d in DEGREES : d>=1}:
    a[d] >= 0 + margin; (17)
```

```
# Intuition: weight for heavy edges >= weight for light edges
subject to HeavyEdge:
    We - Wh <= 0 - margin; (19)
```

```
# collapse parallel edges
subject to parallel {d in DEGREES : d >= 3}:
    Wh - We - 2*a[d] + 2*a[d-1] <= 0 - margin; (22)
```

```
# decomposable edges
subject to Decomposable {d in DEGREES : d >= 1}:
    - a[d] + a[d-1] <= 0 - margin; (25)
```

```
# constraints for the values of h[]
subject to hNotation {d in DEGREES, i in DEGREES : 3 <= i <= d}:
    h[d] - a[i] + a[i-1] <= 0 - margin; (30)(35)
```

```
#####
# constraints for cubic
#####
```

```
# 3-cut
subject to Cut3:
    2*2^(-5*a[3] - 2*h[3]) <= 1 - margin; (31)
```

```
# Independent neighborhood
subject to Indep {q in 0..3}:
    2^(-a[3] - 3*h[3] - q*(Wh-We)) + 2^(-a[3] - 3*h[3] - q*(Wh-We) - 2*(3-q)*h[3])
    <= 1 - margin; (32)
```

```
# One edge in neighborhood
subject to OneEdge1:
    2^(-5*a[3]-h[3]) + 2^(-5*a[3] - 3*h[3]) <= 1 - margin; (34)
```

```
subject to OneEdge2:
    2^(-5*a[3] - h[3] - Wh + We) + 2^(-5*a[3] - h[3] - Wh + We) <= 1 - margin; (33)
```

```
#####
# constraints for degree 4
#####
```

```

# 4-regular

# regular becomes nonregular
subject to Regular41:
    2* 2^(-a[4] - 4*h[4]-R4) <= 1 - margin; (38)

# regular becomes regular
subject to Regular42:
    2* 2^(-5*a[4]) <= 1 - margin; (37)

# 4 non-regular

subject to 4nonregularBase
    {p3p in 0..4, p3pp in 0..4, p4 in 0..3: p3p+p3pp+p4=4}:
    D4n[p3p,p3pp] = -a[4] -(p3p+p3pp)*h[3] -p4*h[4] -ceil(p3pp/2)*(Wh-We); (39)

subject to 4regularBase
    {p3p in 0..4, p3pp in 0..4, p4 in 0..3: p3p+p3pp+p4=4}:
    D4r[p3p,p3pp] = -a[4] -(p3p+p3pp)*a[3] -p4*a[4] -ceil(p3pp/2)*(Wh-We) +R4; (40)

subject to 4nonregularBonus
    {p3p in 0..4, p3pp in 0..4, p4 in 0..3: p3p+p3pp+p4=4}:
    g4n[p3p] = - floor((p3p+2)/3) * (2*h[4]); (41)

subject to 4regularBonus
    {p3p in 0..4, p3pp in 0..4, p4 in 0..3: p3p+p3pp+p4=4}:
    g4r[p3p] = - (floor(p3p/2)+floor(p3p/3)+floor(p3p/4)) * a[3]; (42)

subject to Nonregular41 {p3p in 0..4, p3pp in 0..4, p4 in 0..3: p3p+p3pp+p4=4}:
    2^(D4n[p3p,p3pp]) + 2^(D4n[p3p,p3pp] + g4n[p3p])
    <= 1 - margin; (43)

subject to Nonregular42 {p3p in 0..4, p3pp in 0..4, p4 in 0..3: p3p+p3pp+p4=4}:
    2^(D4n[p3p,p3pp]) + 2^(D4r[p3p,p3pp] + g4r[p3p])
    <= 1 - margin; (44)

subject to Nonregular43 {p3p in 0..4, p3pp in 0..4, p4 in 0..3: p3p+p3pp+p4=4}:
    2^(D4r[p3p,p3pp]) + 2^(D4n[p3p,p3pp] + g4n[p3p])
    <= 1 - margin; (45)

subject to Nonregular44 {p3p in 0..4, p3pp in 0..4, p4 in 0..3: p3p+p3pp+p4=4}:
    2^(D4r[p3p,p3pp]) + 2^(D4r[p3p,p3pp] + g4r[p3p])
    <= 1 - margin; (46)

#####
# constraints for degree 5
#####

# 3-cut for degree 5

```

subject to Cut5_3:
 $2 * 2^{(-a[5] - 6*a[3] + R5 + (Wh-We))} \leq 1 - \text{margin};$ (47)

5-regular

regular becomes nonregular
 subject to Regular51:
 $2 * 2^{(-a[5] - 5*h[5]-R5)} \leq 1 - \text{margin};$ (48)

regular stays regular
 subject to Regular52:
 $2 * 2^{(-6*a[5])} \leq 1 - \text{margin};$ (49)

5 non-regular

clause learning

first branch
 subject to Cf1 {p4 in 1..2, p5 in 3..4: p4+p5=5}:
 $f1[p4] \geq -a[5]-p4*h[4]-p5*h[5];$ (54)

subject to Cf1reg {p4 in 1..2, p5 in 3..4: p4+p5=5}:
 $f1[p4] \geq -a[5]-p4*a[4]-p5*a[5]+R5;$ (54)

second branch, v has at most 1 deg-5 neighbor in N^2
 subject to Cf2a {p4 in 1..2, p5 in 3..4: p4+p5=5}:
 $f2[p4,1] \geq -a[5]-p4*h[4]-p5*h[5]-a[4]-3*h[4]-h[5];$ (55)

subject to Cf2areg {p4 in 1..2, p5 in 3..4: p4+p5=5}:
 $f2[p4,1] \geq -a[5]-p4*a[4]-p5*a[5]+R5;$ (55)

second branch, v (and all other deg-5 nbs of u) has at least 2 deg-5 nbs in N^2
 subject to Cf2b {p4 in 1..2, p5 in 3..4: p4+p5=5}:
 $f2[p4,2] \geq -a[5]-p4*h[4]-p5*h[5]-a[4]-4*h[5];$ (56)

subject to Cf2breg {p4 in 1..2, p5 in 3..4: p4+p5=5}:
 $f2[p4,2] \geq -a[5]-p4*a[4]-p5*a[5]-2*a[3]+R5;$ (56)

additional degree reductions in the 3rd branch (nonregular) for p5=4 vs p5=3
 subject to addDegRedNR53_1:
 $\text{nonreg53} \leq 2*h[5];$ (58)

subject to addDegRedNR53_2:
 $\text{nonreg53} \leq h[4]+h[3]-h[5];$ (59)

third branch, v has at most 1 deg-5 neighbor in N^2
 subject to Cf3a {p4 in 1..2, p5 in 3..4: p4+p5=5}:
 $f3[p4,1] \geq -a[5]-p4*a[4]-p5*a[5]-(4+((4*p4+5*p5-5) \bmod 2))*h[5];$ (57)

subject to Cf3areg {p4 in 1..2, p5 in 3..4: p4+p5=5}:
 $f3[p4,1] \geq -a[5]-p4*a[4]-p5*a[5]-4*a[3]+R5;$ (57)

third branch, v (and all other deg-5 nbs of u) has at least 2 deg-5 nbs in N^2
 subject to Cf3b {p4 in 1..2, p5 in 3..4: p4+p5=5}:

$$f3[p4,2] \geq -a[5]-p4*a[4]-p5*a[5]-6*h[5]-\text{floor}(p5/4)*\text{nonreg53}; \quad (60)$$

subject to Cf3breg {p4 in 1..2, p5 in 3..4: p4+p5=5}:

$$f3[p4,2] \geq -a[5]-p4*a[4]-p5*a[5]-2*a[3]-2*a[5]+R5; \quad (60)$$

the clause learning splitting

subject to Nonregular5cl {p4 in 1..2, nb5 in 1..2}:

$$2^{f1[p4]} + 2^{f2[p4,nb5]} + 2^{f3[p4,nb5]} \leq 1; \quad (61)(62)$$

2-way splitting

2-way splitting, non-reg in both branches, if p3>0, then additional heavy edge

subject to Nonregular51a {p3 in 0..5, p4 in 0..5,
p5 in 0..4, H in 0..1: p3+p4+p5=5
and ((H=1) or (p5 < 3 or p3>0))}:

$$2 * 2^{(-a[5] - p3*h[3] - p4*h[4] - p5*h[5] - H*(Wh-We) - \text{ceil}(p3/5)*(Wh-We))} \leq 1 - \text{margin}; \quad (53)$$

2-way splitting, non-reg in both branches, if p3>0, then additional super-2

subject to Nonregular51b {p3 in 0..5, p4 in 0..5, p5 in 0..4,
H in 0..1: p3+p4+p5=5
and ((H=1) or (p5 < 3 or p3>0))}:

$$2^{(-a[5] - p3*h[3] - p4*h[4] - p5*h[5] - H*(Wh-We) - \text{ceil}(p3/5)*2*h[5])} + 2^{(-a[5] - p3*h[3] - p4*h[4] - p5*h[5] - H*(Wh-We))} \leq 1 - \text{margin}; \quad (52)$$

2-way splitting, becomes reg in both branches

subject to Nonregular52 {p3 in 0..5, p4 in 0..5, p5 in 0..4,
H in 0..1: p3+p4+p5=5
and ((H=1) or (p5 < 3 or p3>0))}:

$$2 * 2^{(-a[5] - p3*a[3] - p4*a[4] - p5*a[5] - H*(Wh-We) + R5)} \leq 1 - \text{margin}; \quad (50)$$

2-way splitting, becomes reg in 1 branch

subject to Nonregular52b {p3 in 0..5, p4 in 0..5, p5 in 0..4,
H in 0..1: p3+p4+p5=5
and ((H=1) or (p5 < 3 or p3>0))}:

$$2^{(-a[5] - p3*a[3] - p4*a[4] - p5*a[5] - H*(Wh-We) + R5)} + 2^{(-a[5] - p3*h[3] - p4*h[4] - p5*h[5] - H*(Wh-We))} \leq 1 - \text{margin}; \quad (50)$$

constraints for degree 6
#####

6-regular

regular becomes nonregular

subject to Regular61:

$$2 * 2^{(-a[6] - 6*h[6]-R6)} \leq 1 - \text{margin}; \quad (64)$$

regular stays regular
subject to Regular62:
 $2 * 2^{(-7*a[6])} \leq 1 - \text{margin};$ (63)

6 non-regular

nonregular stays nonregular
subject to Nonregular61 {p3 in 0..6, p4 in 0..6, p5 in 0..6, p6 in 0..5:
 $p3+p4+p5+p6=6$):
 $2 * 2^{(-a[6] - p6*h[6] - p5*h[5] - p4*h[4] - p3*h[3])} \leq 1 - \text{margin};$ (65)

nonregular becomes regular
subject to Nonregular62 {p3 in 0..6, p4 in 0..6, p5 in 0..6, p6 in 0..5:
 $p3+p4+p5+p6=6$):
 $2 * 2^{(-a[6] - p6*a[6] - p5*a[5] - p4*a[4] - p3*a[3] +R6)} \leq 1 - \text{margin};$ (66)

(Serge Gaspers) INSTITUTE OF INFORMATION SYSTEMS (184/3), VIENNA UNIVERSITY OF TECHNOLOGY, FAVORITENSTRASSE 9-11, A-1040 VIENNA, AUSTRIA
E-mail address: gaspers@kr.tuwien.ac.at

(Gregory B. Sorkin) DEPARTMENT OF MANAGEMENT, LONDON SCHOOL OF ECONOMICS, HOUGHTON STREET, LONDON WC2A 2AE, UK
E-mail address: g.b.sorkin@lse.ac.uk