PRIFYSGOL CYMRU ABERTAWE
UNIVERSITY OF WALES SWANSEA

UNIVERSITY OF WALES SWANSEA

REPORT SERIES

Polynomial time SAT decision for linearly lean complement-invariant clause-sets of minimal reduced deficiency

by

Oliver Kullmann

Report # CSR 1-2007

Computer Science
Gwyddor Cyfrifiadur

Polynomial time SAT decision for linearly lean complement-invariant clause-sets of minimal reduced deficiency

Oliver Kullmann[1,⋆]

Computer Science Department, University of Swansea
Swansea, SA2 8PP, UK
O.Kullmann@Swansea.ac.uk, http://cs-svr1.swan.ac.uk/~csoliver

Abstract. We study *complement-invariant clause-sets* F, where for every clause $C \in F$ we have $\overline{C} = \{\overline{x} : x \in C\} \in F$, i.e., F is closed under elementwise complementation of clauses. The *reduced deficiency* of a clause-set F is defined as $\delta_{\mathrm{r}}(F) := \frac{1}{2}(\delta(F) - n(F))$, where $\delta(F) = c(F) - n(F)$ is the difference of the number of clauses and the number of variables, while the maximal reduced deficiency is $\delta^*_{\mathrm{r}}(F) := \max_{F' \subseteq F} \delta_{\mathrm{r}}(F') \geq 0$. If F is complement-invariant and linearly lean (has no non-trivial linear autarkies) then we have $\delta^*_{\mathrm{r}}(F) = \delta_{\mathrm{r}}(F)$. We show polynomial time SAT decision for complement-invariant clause-sets F with $\delta^*_{\mathrm{r}}(F) = 0$, exploiting the (non-trivial) decision algorithm for sign-non-singular matrices given by [Robertson, Seymour, Thomas 1999]. Minimally unsatisfiable complement-invariant clause-sets F fulfil $\delta_{\mathrm{r}}(F) \geq 0$, and we immediately get polynomial time decidability of minimally unsatisfiable complement-invariant clause-sets F with $\delta_{\mathrm{r}}(F) = 0$, but we can give also somewhat more direct algorithms and characterisations (especially for sub-classes).

1 Introduction

The *deficiency* $\delta(F) = c(F) - n(F)$ of clause-sets (where $c(F)$ is the number of clauses, and $n(F)$ is the number of variables) is an interesting parameter, allowing polynomial time SAT decision for clause-sets F with bounded maximal deficiency (maximised over all sub-clause-sets). We make the first step towards an analogous poly-time hierarchy, where we consider only *complement-invariant* clause-sets F (for every $C \in F$ also $\overline{C} \in F$ holds), while we strengthen the notion of deficiency to *reduced deficiency* $\delta_{\mathrm{r}}(F) = \frac{1}{2}(\delta(F) - n(F))$. We proof polynomial time SAT decision for complement-invariant clause-sets with the maximal reduced deficiency equal to its minimal value 0. The proof is based on a recent breakthrough by [10], where several long outstanding problems (open for up to 93 years) have been solved. As an application we obtain polynomial time decision of 2-colourability of hypergraphs with the maximal deficiency equal to

⋆ Supported by grant EPSRC GR/S58393/01.

its minimal value 0 (where the deficiency of an hypergraph is the difference of the number of hyperedges and the number of variables).

We exploit heavily the relationship between autarky theory and matrix analysis. More specifically, we exploit *qualitative matrix analysis* (QMA), matrix analysis modulo the equivalence relation between matrices given by having the same sign pattern. [2] describes the foundations of QMA in qualitative economics:

> Qualitative economics is usually considered to have originated with the work of Samuelson who discussed the possibility of determining unambiguously the qualitative behavior of solution values of a system of equations. In his pioneering paper Lancaster put it this way:
> Economists believed for a very long time, and most economists would still hope it to be so, that a considerable body of sensible economic propositions could be expressed in a qualitative way, that is, in a form in which the algebraic sign of some effect is predicted from a knowledge of the signs, only, of the relevant structural parameters of the system.

For example, that a system of (differential) equations has a unique solution is (often) controlled by a square matrix A having non-zero determinant. Now in qualitative matrix analysis we want $\det(A) \neq 0$ independent of the magnitude of the entries of A, if only the signs are preserved. This example leads to two fundamental notions of QMA, which are also central for this article: A square matrix A is called an **SNS-matrix** if all matrices A' with the same sign pattern are non-singular, while more generally a matrix A is called an **L-matrix** if all matrices A' with the same sign pattern have linearly independent rows. As already remarked in [5], L-matrices correspond 1-1 to complement invariant clause-sets which are lean (have no non-trivial autarky); the main results of this article are related to SNS-matrices which correspond 1-1 to lean complement invariant clause-sets with reduced deficiency 0.

Likely the best way to determine whether an arbitrary matrix is an L-matrix is to determine whether the corresponding complement invariant clause-set is lean (for example using the method as discussed in [7]). However for SNS-matrices finally a polynomial time algorithm was found in [10], which is the basis for deciding leanness and minimally unsatisfiability for complement invariant clause-sets with reduced deficiency 0 in polynomial time (Corollary 14). In Theorem 20 these result are strengthened by showing that actually satisfiability is decidable in polynomial time for complement invariant clause-sets with *maximal* reduced deficiency 0. As an immediate application, in Corollary 22 we obtain, that 2-colourability for hypergraphs G with maximal deficiency 0 is decidable in polynomial time (where the deficiency is the difference of the number of hyperedges and the number of vertices).

2 Some general theory of autarky systems

We use standard (boolean) clause-sets F, which are finite sets of clauses here, where a clause is a finite set of non-clashing literals; the empty clause is $\bot$, the empty clause-set $\top$. Application of partial assignments φ to F is denoted by $\varphi * F$ (substituting truth values for the literals touched by φ with subsequent simplification), while by $V * F$ for some set V of variables the operation of crossing out the variables of V from F is denoted (that is, $V * F = \{\{x \in C : \mathrm{var}(x) \notin V\} : C \in F\}$). Furthermore $F[V] := ((\mathrm{var}(F) \setminus V) * F) \setminus \{\bot\}$ is the restriction of F to the variable-set V.

A partial assignment φ is an **autarky** for a clause-set F if every clause $C \in F$ touched by φ (that is, $\mathrm{var}(\varphi) \cap \mathrm{var}(C) \neq \emptyset$) is actually satisfied by φ; the systematic exploration of autarkies started with [3], where the reader may find more background information. The main algorithmic use of autarkies φ for F is *autarky reduction*, the transition from F to the satisfiability-equivalent $\varphi * F$ (which is the sub-clause-set of F given by all clauses not touched by φ). An autarky φ for F is called **balanced** if also $\overline{\varphi}$ is an autarky for F, where $\overline{\varphi}$ is the pointwise complement of φ. Balanced autarkies share many general properties with ordinary autarkies. Since there are many more types of "special autarkies", a general theory is needed, provided by the theory of **autarky systems**, which axiomatically specifies the properties needed of the map $F \mapsto \mathcal{A}(F) \subseteq \mathrm{Auk}(F)$ from clause-sets F to sub-monoids $\mathcal{A}(F)$ of the autarky monoid so that the special autarkies behave like ordinary autarkies. See [4,5,6] for precise definitions and some fundamental results. All autarky systems we have encountered fulfil, besides the basic requirements cast in the notion of an autarky system, some other fundamental properties (or can be easily extended to fulfil them), which are collected in the notion of a **normal autarky system** (see [6] for the most current version of this notion):

1. $\mathcal{A}$ is **iterative**, if $\mathcal{A}$-autarkies of sub-clause-sets of F obtained by $\mathcal{A}$-autarky reduction are still in $\mathcal{A}(F)$.
2. $\mathcal{A}$ is called **standardised**, if $\mathcal{A}$-autarkies can be set arbitrarily on variables not in F.
3. $\mathcal{A}$ is **$\bot$-invariant**, if $\mathcal{A}(F)$ is invariant against addition or deletion of the empty clause $\bot$.
4. $\mathcal{A}$ is **stable under variable elimination**, if for every set V of variables the $\mathcal{A}$-autarkies of $V * F$ which do not use V are exactly the $\mathcal{A}$-autarkies of F which do not use V.
5. $\mathcal{A}$ is **invariant under renaming**, if renaming of clause-sets F carries over to $\mathcal{A}(F)$.

In this article we consider four normal autarky systems:[1)]

[1)] The problem with simple (balanced) linear autarkies, which makes "normalisation" necessary, is that these autarky systems are not "iterative", that is, in general it is not the case that if φ is a simple (balanced) linear autarky for F and if ψ is a simple (balanced) linear autarky for $\varphi * F$, then $\psi \circ \varphi$ is a (balanced) linear autarky for F, because ψ might invalidate the special conditions on these autarkies for clauses already satisfied by φ. "Normalisation" just adds these compositions of "iterated autarkies" to the autarky monoid, and for the resulting (balanced) linear autarkies the adjective "simple" is dropped; see Subsections 4.4 and 4.6 in [3] for the special case of linear autarkies, and Lemma 8.4 in [5] for the general case.

– Auk, the full autarky system (all autarkies)
– BAuk, the system of balanced autarkies (see Section 3)
– LAuk, the normalised system created from simple linear autarkies (see Section 4)
– BLAuk, the normalised system created from simple balanced linear autarkies (see Section 4).

Let $\mathcal{A}$ be a normal autarky system. A clause-set F is called **$\mathcal{A}$-lean** if F has no non-trivial $\mathcal{A}$-autarky (one which touches F); there is a largest $\mathcal{A}$-lean sub-clause-set of F, called the **$\mathcal{A}$-lean kernel**. F is called **$\mathcal{A}$-satisfiable** if the $\mathcal{A}$-lean kernel is $\top$ (which is equivalent to the existence of an $\mathcal{A}$-autarky for F which actually satisfies F), while otherwise F is called **$\mathcal{A}$-unsatisfiable**. $\top$ is $\mathcal{A}$-lean, and if $F \neq \top$ is $\mathcal{A}$-lean, then F is $\mathcal{A}$-unsatisfiable; F is $\mathcal{A}$-lean iff $F \setminus \{\bot\}$ is $\mathcal{A}$-lean iff $F \cup \{\bot\}$ is $\mathcal{A}$-lean. If $\mathcal{A}$ is the full autarky system, then we just speak of satisfiability, unsatisfiability, and leanness.

Definition 1. *A clause-set F is called **minimally $\mathcal{A}$-unsatisfiable**, if F is $\mathcal{A}$-unsatisfiable, while every $F' \subset F$ is $\mathcal{A}$-satisfiable.*

1. If $\mathcal{A}$ is the full autarky system, then we obtain ordinary minimal unsatisfiability.
2. $F \neq \top$ is minimally $\mathcal{A}$-lean (that is, F is $\mathcal{A}$-lean, while every $F' \subset F$ is not $\mathcal{A}$-lean) if and only if F is minimally $\mathcal{A}$-unsatisfiable.

Definition 2. *A clause-set F is called **barely $\mathcal{A}$-lean** (in generalisation of the notion of a "barely L-matrix" in [2]) if F is $\mathcal{A}$-lean, while for every clause $C \in F$ the clause-set $F \setminus \{C\}$ is not $\mathcal{A}$-lean.*

1. $\top$ is barely $\mathcal{A}$-lean, while if $c(F) = 1$, then F is not barely $\mathcal{A}$-lean.
2. If F is minimally $\mathcal{A}$-unsatisfiable and $c(F) \geq 2$, then F is barely $\mathcal{A}$-lean.
3. If $\mathcal{A}$ is the full autarky system, then we speak of "barely lean".
4. F is barely $\mathcal{A}$-lean iff $F \setminus \{\bot\}$ is barely $\mathcal{A}$-lean iff $F \cup \{\bot\}$ is barely $\mathcal{A}$-lean.

Lemma 3. *If for a normal autarky system $\mathcal{A}$ the autarky-existence problem (deciding whether a non-trivial autarky exists, and finding one if existent) is solvable in polynomial time, then the following decision problems are in polynomial time:*

1. *$\mathcal{A}$-satisfiability and $\mathcal{A}$-unsatisfiability;*
2. *minimal $\mathcal{A}$-unsatisfiability;*
3. *A-leanness;*
4. *barely $\mathcal{A}$-leanness.*

Definition 4. *A clause-set F is a **generalised sum** of clause-sets F_1, F_2 if $var(F_1) \cap \mathrm{var}(F_2) = \emptyset$ with $\mathrm{var}(F_1), \mathrm{var}(F_2) \neq \emptyset$, and there is F_2' with $F = F_1 \mathbin{\dot\cup} F_2'$, such that F_2' is obtained from F_2 by adding literals over $\mathrm{var}(F_1)$ to clauses from F_2 (that is, there exists a bijection $\alpha : F_2 \to F_2'$ such that for all $C \in F_2$ we have $\alpha(C) \supseteq C$ and $\mathrm{var}(\alpha(C) \setminus C) \subseteq \mathrm{var}(F_1)$).*

1. Let F be a generalised sum of F_1, F_2:
   (a) $F = F_1 \mathbin{\dot\cup} F_2'$ and $\mathrm{var}(F) = \mathrm{var}(F_1) \mathbin{\dot\cup} \mathrm{var}(F_2)$, and thus $\delta(F) = c(F) - n(F) = (c(F_1) + c(F_2)) - (n(F_1) + n(F_2)) = \delta(F_1) + \delta(F_2)$.
   (b) $F_2 \setminus \{\bot\} = F[\mathrm{var}(F_2)]$.
   (c) If F is $\mathcal{A}$-lean, then so is F_2.
   (d) If F_1, F_2 are $\mathcal{A}$-lean, then so is F.
2. If F is any clause-set and $F_1 \subset F$ with $\emptyset \subset \mathrm{var}(F_1) \subset \mathrm{var}(F)$ and such that for $C \in F \setminus F_1$ we have $\mathrm{var}(C) \not\subseteq \mathrm{var}(F_1)$, then F is a generalised sum of F_1, F_2 for $F_2 := F[\mathrm{var}(F) \setminus \mathrm{var}(F_1)]$.

Definition 5. *A clause-set F is **$\mathcal{A}$-indecomposable** (in generalisation of the notion of an "L-indecomposable matrix" in [2]) if F is not the generalised sum of $\mathcal{A}$-lean clause-sets F_1, F_2; otherwise F is called **$\mathcal{A}$-decomposable**.*

1. Note that if F is $\mathcal{A}$-decomposable, then F is necessarily $\mathcal{A}$-lean, while if F is not $\mathcal{A}$-lean, then F is $\mathcal{A}$-indecomposable.
2. If $\mathcal{A}$ is the full autarky system, then we speak of "autarky indecomposable".
3. F is $\mathcal{A}$-decomposable iff $F \setminus \{\bot\}$ is $\mathcal{A}$-decomposable iff $F \cup \{\bot\}$ is $\mathcal{A}$-decomposable.

Generalising Theorem 2.2.5 in [2]:

Lemma 6. *A clause-set F with $c(F) \geq 2$ is minimally $\mathcal{A}$-unsatisfiable if and only if the following two conditions hold:*

(i) F is barely $\mathcal{A}$-lean
(ii) F is $\mathcal{A}$-indecomposable.

Proof. Clearly the two conditions are necessary; it remains to see that they are sufficient. Since F is barely $\mathcal{A}$-lean, F is $\mathcal{A}$-lean, and thus $\mathcal{A}$-unsatisfiable. We have $\bot \notin F$, since otherwise $F \setminus \{\bot\}$ would not be $\mathcal{A}$-satisfiable. Now consider $C \in F$, and assume that $F \setminus \{C\}$ is not $\mathcal{A}$-satisfiable. Then there is a non-trivial autarky φ for $F \setminus \{C\}$ such that $F_1 := \varphi * (F \setminus \{C\})$ is $\mathcal{A}$-lean, where $\top \subset F_1 \subset F \setminus \{C\}$. Now, by Remark 2 to Definition 4, F is a generalised sum of F_1 and $F_2 := F[\mathrm{var}(F) \setminus \mathrm{var}(F_1)]$ (note that φ touches C while not satisfying C) contradicting $\mathcal{A}$-indecomposability of F. □

We conclude this section on general autarky systems by regarding the complexity of the basic decision problems for the full autarky system:

1. satisfiability/unsatisfiability decision is NP/coNP-complete;
2. minimally unsatisfiability decision is D^P-complete ([9]);
3. leanness decision is coNP-complete ([5]);
4. barely lean decision is D^P-complete ([8]);
5. autarky decomposability decision is in Σ_2 (while more is not known; possibly it is Σ_2-complete).

3 Balanced autarkies

[5] introduced **balanced autarkies** for clause-sets F, which are partial assignments φ such that for every clause $C \in F$ touched by φ there exists a satisfied *as well as* a falsified literal in C. The set of all balanced autarkies for F is $\mathbf{BAuk}(\boldsymbol{F})$; it is BAuk a normal autarky system. We use the following phrases:

- "BAuk-satisfiable" resp. "BAuk-unsatisfiable" is called *balanced satisfiable* resp. *balanced unsatisfiable*;
- "minimally BAuk-unsatisfiable" is called *minimally balanced unsatisfiable*;
- "(barely) BAuk-lean" is called *(barely) balanced lean*;
- "BAuk-indecomposable" is called *balanced autarky-indecomposable*.

The complexities of the basic decision problems for balanced autarkies are likely the same as for general autarkies (see the end of Section 2), but proven at this time is only the coNP-completeness of balanced leanness decision (as remarked in [5]).

For a partial assignment by $\overline{\boldsymbol{\varphi}}$ we denote the pointwise complement of φ, that is, $\mathrm{var}(\overline{\varphi}) = \mathrm{var}(\varphi)$ and $\overline{\varphi}(v) = \overline{\varphi(v)}$.

Lemma 7. *The following assertions are equivalent for a partial assignment* φ *and a clause-set* F*:*

1. φ *is a balanced autarky for* F
2. φ *and* $\overline{\varphi}$ *are autarkies for* F*.*

Thus complementation of partial assignments yields an automorphism of the balanced autarky monoid $\mathrm{BAuk}(F)$*.*

3.1 *L*-matrices and SNS-matrices

All matrices in this article have real entries. For a matrix M the *sign pattern* $\mathrm{sgn}(M)$ is the $\{-1, 0, +1\}$-matrix $\mathrm{sgn}(M)$ of the same dimension given by entrywise sgn-formation, while the *null pattern* of M is $\mathrm{sgn}(|M|)$ (a $\{0, 1\}$-matrix), where $|M|$ denotes entrywise absolute-value formation.

For a clause-set F let $M(F)$ be the *clause-variable matrix* of F (see Section 3 in [5] for more details). As shown in Section 5 of [5], a clause-set F is balanced lean if and only if the matrix $M(F)^{\mathrm{t}}$ is an *L-matrix*, where a matrix M is called an L-matrix if each matrix with the same sign pattern as M has linearly independent rows. L matrices have at least as many columns as rows, and thus for balanced lean clause-sets F we have $\delta(F) \geq 0$ (see Lemma 19 for a more general and stronger statement). Square L-matrices are called *SNS-matrices* ("sign-non-singular matrices"), which are characterised by the condition that every square matrix with the same sign pattern is invertible (non-singular); so a clause-set F with $\delta(F) = 0$ is balanced lean iff $M(F)^{\mathrm{t}}$ is an SNS-matrix. We also have the inverse directions (so that the notions of L- and SNS-matrices are fully captured by balanced lean clause-sets):

1. A matrix M without repeated columns is an L-matrix if and only if M has no zero rows and a clause-set F with $M(F)^{t} = M$ is balanced lean.
2. A matrix M is an SNS-matrix if and only if M neither has zero rows nor repeated columns and a clause-set F with $M(F)^{t} = M$ is balanced lean and fulfils $\delta(F) = 0$.

The special treatment of rows and columns is necessary due to the use of clause-sets, which contract multiple clauses and eliminate purely "formal" variables (which do not occur).

For a square $\{-1, 0, +1\}$-matrix A of order $n \in \mathbb{N}_0$ the following conditions are equivalent (see [2] for the (easy) proofs):

1. A is an SNS-matrix;
2. $\det(A) \neq 0$ and all non-null terms in the determinant expansion of A have the same sign (that is, there is $\varepsilon \in \{-1, +1\}$ such for all permutations $\pi \in S_n$ of $\{1, \dots, n\}$ in case of $\prod_{i=1}^{n} A_{i,\pi(i)} \neq 0$ we have $\operatorname{sgn}(\pi) \cdot \prod_{i=1}^{n} A_{i,\pi(i)} = \varepsilon$).
3. $\det(A) \neq 0$ and $\operatorname{per}(|A|) = |\det(A)|$, where $\operatorname{per}(A) = \sum_{\pi \in S_n} \prod_{i=1}^{n} A_{i,\pi(i)}$ denotes the permanent of a square matrix A.
4. $\det(A) \neq 0$ and for every square matrix M of order n we have $\operatorname{per}(M * |A|) = |\det(M * A)|$, where $M * A$ denotes the pointwise (Hadamard-)product of matrices of the same dimension.

The "Pólya-Problem" as discussed in [10] is the problem to determine, whether for a square $\{0, 1\}$-matrix A there exists an SNS-$\{-1, 0, +1\}$-matrix B with the same null pattern. Algorithm 9.7 there decides in polynomial time whether for input A the matrix B exists, and also computes B if it exists (while deciding the L-matrix-property is coNP-complete).2) This algorithm yields poly-time decision of the SNS-property for square $\{-1, 0, +1\}$-matrices M as follows:

1. If $\det(M) = 0$ then M is not an SNS-matrix.
2. Let $A := |M|$.
3. If A has no associated B then M is not an SNS-matrix.
4. Otherwise M is an SNS-matrix iff $|\det(B)| = |\det(M)|$.

2) [10] uses a (directly) equivalent formulation based on the notion of "Pfaffian orientation". Given a bipartite graph G with a bipartition, where both parts are of the same size, the permanent of the "reduced adjacency matrix" of G (a square $\{0, 1\}$-matrix indicating the connections between both sides) is the number of perfect matchings of G; more generally for an arbitrary graph G, the Pfaffian of the (skew-symmetric) matrix S obtained by putting the adjacency matrix $A(G)$ in the upper right corner and $-A(G)$ in the lower left corner (while the rest is zero) is the number of perfect matchings (or 1-factors) of G. A Pfaffian orientation of G is an orientation of the edges such that the absolute value of the determinant of (the skew-symmetric) S', obtained from S by switching signs in S accordingly, yields the number of perfect matchings of G. When applied to bipartite graphs, the problem of deciding whether a Pfaffian orientation of G exists is just the Pólya-problem.

A problem remains: In this way we can decide whether M is an SNS-matrix, but in case M is not an SNS-matrix, how do we find (in polynomial time) a matrix M' with the same sign pattern as M which is singular ?! The critical step is Step 3 in the above procedure, where one has to examine the different obstructions studied in [10]. It seems plausible, that from these obstructions one can compute a witness M', but one the other hand it doesn't seem to be straight-forward, and in order to be careful, we formulate the remaining algorithmic problem as a conjecture:

Conjecture 8. The following functional computation problem can be solved in polynomial time: Given a square matrix A over $\{-1, 0, +1\}$, if A is not an SNS-matrix, then a matrix A' over $\mathbb{Q}$ with the same sign pattern as A' can be computed such that A' is singular.

3.2 Complement-invariant clause-sets

By $\boldsymbol{\overline{F}} := \{\overline{C} : C \in F\}$ we denote the clause-wise complement of a clause-set F, where $\boldsymbol{\overline{C}} := \{\overline{x} : x \in C\}$). The following lemma shows that (un)satisfiability, minimal unsatisfiability and leanness of $F \cup \overline{F}$ is equivalent to the respective property for F where the full autarky system is replaced with the autarky system of balanced autarkies (other properties behave in a more complicated way; we consider only barely leanness here).

Lemma 9. *For a clause-set F we have:*

1. *For a partial assignment φ the following conditions are equivalent:*
   (a) *φ is an autarky for $F \cup \overline{F}$.*
   (b) *φ is a balanced autarky for $F \cup \overline{F}$.*
   (c) *φ is a balanced autarky for F.*
2. *F is balanced satisfiable resp. balanced unsatisfiable iff $F \cup \overline{F}$ is satisfiable resp. unsatisfiable.*
3. *F is minimally balanced unsatisfiable iff $F \cup \overline{F}$ is minimally unsatisfiable.*
4. *F is balanced lean iff $F \cup \overline{F}$ is lean.*
5. (a) *If F is barely balanced lean then $F \cup \overline{F}$ is barely lean.*
   (b) *Conversely, assume that $F \cup \overline{F}$ is barely lean.*
      i. *If $c(F) = 1$ (here the clause of F must be a unit clause), then F is not barely balanced lean.*
      ii. *If F is a generalised sum of some $F_1, \{U\}$, where U is a unit clause, then F is not barely balanced lean (since F_1 is balanced lean).*
      iii. *Otherwise F is barely balanced lean.*

Proof. Part 1 follows by definition, and Parts 2, 4 are direct consequences of this basic fact. For Part 3 it is left to show that if F is minimally balanced unsatisfiable then $F \cup \overline{F}$ is minimally unsatisfiable; consider a clause $C \in F \cup \overline{F}$, and we have to show that $(F \cup \overline{F}) \setminus \{C\}$ is satisfiable. There is a balanced satisfying assignment φ for $(F \cup \overline{F}) \setminus \{C, \overline{C}\}$. If $C = \overline{C}$ ($\Leftrightarrow C = \bot$), then we are done. If $|C| = 1$, then $F = \{C\}$, and we are done as well. So assume $|C| \geq 2$. Now φ touches C (since F is balanced unsatisfiable), and then φ or $\overline{\varphi}$ is a satisfying assignment for $F \cup \overline{F}$. For Part 5 first assume that F is barely balanced lean, and we have to show that $F \cup \overline{F}$ is barely lean. Consider $C \in F \cup \overline{F}$. Now $(F \cup \overline{F}) \setminus \{C, \overline{C}\}$ has a non-trivial balanced autarky φ. Since $F \cup \overline{F}$ is lean, φ touches C, and then φ or $\overline{\varphi}$ is a non-trivial autarky for $(F \cup \overline{F}) \setminus \{C\}$. Now assume that $F \cup \overline{F}$ is barely lean, $c(F) \geq 2$, and consider $C \in F$. There is a non-trivial autarky φ for $(F \cup \overline{F}) \setminus \{C\}$. If φ touches $F \setminus \{C\}$, then φ is a non-trivial balanced autarky for $F \setminus \{C\}$. Otherwise there is be $x \in C$ with $\mathrm{var}(x) \notin \mathrm{var}(F \setminus \{C\})$ (while $\mathrm{var}(C \setminus \{x\}) \subseteq \mathrm{var}(F \setminus \{C\}$, since F is lean), and thus F is a generalised sum of $F \setminus \{C\}, \{x\}$ (note that in case of $\mathrm{var}(F \setminus \{C\}) = \emptyset$ we would have $F \setminus \{C\} = \{\bot\}$, and then $F \cup \overline{F}$ would not be lean). Now $F \setminus \{C\}$ is balanced lean (a non-trivial balanced autarky for $F \setminus \{C\}$ either does not touch C, or otherwise a non-balancedness can be repaired using x, and so in both cases we contradict that F is balanced lean), and thus F is not barely balanced lean. □

By Lemma 9, Part 3 together with Lemma 6 we get:

Corollary 10. *A clause-set $F \cup \overline{F}$, where F is a clause-set with $c(F) \geq 2$, is minimally unsatisfiable if and only if F is barely balanced lean and balanced autarky-indecomposable.*

Definition 11. *A clause-set F is called **complement-invariant** if $F = \overline{F}$, which is equivalent to the existence of a clause-set F_0 with $F = F_0 \cup \overline{F_0}$; such an F_0 is called a **core half** of F if $2 \cdot c(F_0) = c(F)$.*

So only complement-invariant clause-sets F with $\bot \notin F$ have a core half (which is unique only up to complement of the clauses), but this little inconvenience seems not to justify the use of multi-clause-sets instead in this article (the problem is that $\overline{\bot} = \bot$, which causes contraction for clause-sets). If F is complement-invariant and φ is an autarky for F, then also $\varphi * F$ is complement-invariant.

Definition 12. *For an arbitrary clause-set F we define the **reduced deficiency** $\boldsymbol{\delta_{\mathrm{r}}(F)} := \frac{1}{2}(\delta(F) - n(F)) \in \frac{1}{2}\mathbb{N}_0$.*

If F is complement-invariant with core half F_0, then we have $\delta_{\mathrm{r}}(F) = \delta(F_0)$, since

$$\delta_{\mathrm{r}}(F) = \frac{1}{2}(\delta(F) - n(F)) = \frac{1}{2}(c(F) - 2n(F)) = \frac{1}{2}(2c(F_0) - 2n(F)) = \\ c(F_0) - n(F) = c(F_0) - n(F_0) = \delta(F_0).$$

If F is lean, then we have $\delta_{\mathrm{r}}(F) = \delta(F_0) \geq 0$ (see Lemma 19 for a more general statement).

3.3 Square balanced lean clause-sets

Theorem 13. *Consider a clause-set F with $\delta(F) = 0$.*

1. *It is decidable in poly-time whether F is balanced lean.*
2. *It is decidable in poly-time whether F is barely balanced lean.*
3. *Assume F is balanced lean. Then F is balanced autarky-decomposable if and only if F is a generalised sum of clause-sets F_1, F_2 with $\delta(F_1) = \delta(F_2) = 0$.*
4. *It is decidable in poly-time whether F is balanced autarky-indecomposable.*
5. *It is decidable in poly-time whether F is minimally balanced unsatisfiable.*

Proof. Part 1 follows by [10] as discussed in Subsection 3.1. For Part 2 first we test whether F is balanced lean (by Part 1); assume now that F is balanced lean, and consider $F \in C$. We have to test whether $F \setminus \{C\}$ is (not) balanced lean. If $\delta(F \setminus \{C\}) = 0$, then we can apply Part 1, and so assume $\delta(F \setminus \{C\}) \neq 0$. If C would contain two or more variables not occurring in $F \setminus \{C\}$ then F would not be balanced lean, and so all variables of C must occur in $F \setminus \{C\}$, and we have $\delta(F \setminus \{C\}) = -1$, in which case $F \setminus \{C\}$ is not balanced lean (as desired). The statement of Part 3 is essentially equivalent to Theorem 2.2.1 in [2], however for the sake of completeness, and also since our notions differ slightly from [2] in order to accommodate for the differences in handling matrices and clause-sets, we give a proof here. If F is balanced autarky-decomposable, then F is a generalised sum of balanced lean clause-sets F_1, F_2; we then have $\delta(F_1), \delta(F_2) \geq 0$, and due to $\delta(F) = \delta(F_1) + \delta(F_2)$ we get $\delta(F_1) = \delta(F_2) = 0$. For the opposite direction assume that F is a generalised sum of clause-sets F_1, F_2 with $\delta(F_1) = \delta(F_2) = 0$; we have that F_2 is balanced lean, and we show that F_1 is also balanced lean, which is equivalent to $M(F_1)$ being an SNS-matrix. We have the matrix decomposition

$$M(F) = \begin{pmatrix} M(F_1) & 0 \\ * & M(F_2) \end{pmatrix}.$$

Because of $\det(M(F)) = \det(M(F_1)) \cdot \det(M(F_2))$ we get $\det(M_1) \neq 0$, and moreover, if there would be two non-null terms in the determinant expansion of $M(F_1)$ then we would get two non-null terms in the determinant expansion of $M(F)$ contradicting that $M(F)$ is an SNS-matrix (see the characterisation of SNS-matrices in Subsection 3.1). For Part 4 w.l.o.g. we can assume that $\bot \notin F$; it suffices now to realise that F is a generalised sum of clause-sets F_1, F_2 with $\delta(F_1) = \delta(F_2) = 0$ iff $M(F)$ is partly decomposable as defined in Subsection 4.2 of [1] (there the square submatrices in the decomposition are allowed to be zero matrices, however this cannot happen in our case, since $M(F)$ has no zero column as a clause-variable matrix, while it has no zero row by assumption), where the property of being partly decomposable is decidable in polynomial time.[3)] Finally Part 5 follows with Lemma 6 and Parts 2, 4 (and using that $\top$ is not minimally balanced unsatisfiable while $\{C\}$ for some unit clause C is minimally balanced unsatisfiable). □

[3)] By Corollary 4.2.4 in [1] thus the algorithm for deciding whether F is balanced autarky-indecomposable works as follows: Let $A := |M(F)|$, and consider the bipar-

Corollary 14. *Consider a complement-invariant clause-set F with $\delta_r(F) = 0$.*

1. *It is decidable in poly-time whether F is lean.*
2. *It is decidable in poly-time whether F is barely lean.*
3. *It is decidable in poly-time whether F is minimally unsatisfiable.*

Proof. Part 1 follows with Theorem 13, Part 1 and Lemma 9, Part 4. Part 2 follows with Theorem 13, Part 2 and Lemma 9, Part 5 as follows: W.l.o.g. $\bot \notin F$. Let F_0 be a core half of F. If F_0 is not balanced lean, then F is not barely lean; assume now that F_0 is balanced lean. If F_0 is barely balanced lean, so is F; assume now that F_0 is not barely balanced lean. If $F_0 = \{C\}$, then F is barely lean iff C is a unit clause. If F_0 is not a generalised sum of $F'_0, \{C\}$ for some unit clause C, then F is not barely lean. The remaining case is that F_0 is a generalised sum of $F'_0, \{C\}$ for some unit clause C. Since F_0 is balanced lean, also F'_0 must be balanced lean here; now F is barely lean iff F'_0 is barely balanced lean. Finally Part 3 follows with Theorem 13, Part 5 and Lemma 9, Part 3. □

4 Linear autarkies and balanced linear autarkies

A *linear autarky* for a clause-set F is given by a non-trivial solution to the linear-programming problem $M(F) \cdot x \geq 0$, while a *balanced linear autarky* for F is given by a non-trivial solution to the linear algebra problem $M(F) \cdot x = 0$; see [3,5] for more information. Balanced linear autarkies are special balanced autarkies, and a partial assignment φ is a balanced linear autarky for F iff $\varphi, \overline{\varphi}$ are linear autarkies for F. The autarky-existence problem is solvable in polynomial time for linear autarkies as well as for balanced linear autarkies, and thus with Lemma 3 the basic decision problems for these two (normal) autarky systems are solvable in polynomial time (while it remains to determine the complexity of (balanced) linear-autarky-decomposability decision). In full analogy to Lemma 9 we have, replacing "lean" by "linearly lean" in the statements as well as in the proofs:

Lemma 15. *For a clause-set F we have:*

1. *For a partial assignment φ the following conditions are equivalent:*
   (a) φ is a linear autarky for $F \cup \overline{F}$.
   (b) φ is a balanced linear autarky for $F \cup \overline{F}$.
   (c) φ is a balanced linear autarky for F.
2. *F is balanced linearly satisfiable resp. balanced linearly unsatisfiable iff $F \cup \overline{F}$ is linearly satisfiable resp. linearly unsatisfiable.*

---

tite graph G with reduced adjacency matrix A. If G has no perfect matching then F is balanced autarky-decomposable; otherwise obtain A' from A by permuting accordingly rows and columns of A in such a way that the main diagonal of A' has only entries equal to 1. Now F is balanced autarky-indecomposable iff the directed graph with adjacency matrix A' is strongly connected.

3. *F is minimally balanced linearly unsatisfiable iff* $F \cup \overline{F}$ *is minimally linearly unsatisfiable.*
4. *F is balanced linearly lean iff* $F \cup \overline{F}$ *is linearly lean.*
5. *(a) If F is barely balanced linearly lean then* $F \cup \overline{F}$ *is barely linearly lean.*
   (b) Conversely, assume that $F \cup \overline{F}$ *is barely linearly lean.*
      i. If $c(F) = 1$ *(here the clause of F must be a unit clause), then F is not barely balanced linearly lean.*
      ii. If F is a generalised sum of some $F_1, \{U\}$*, where U is a unit clause, then F is not barely balanced linearly lean (since* F_1 *is balanced lean).*
      iii. Otherwise F is barely balanced linearly lean.

By Lemma 6 we get (analogously to Corollary 10):

Corollary 16. *A clause-set* $F \cup \overline{F}$ *for* $c(F) \geq 2$ *is minimally linearly unsatisfiable if and only if F is barely balanced linearly lean and balanced linear-autarky-indecomposable.*

Analogously to Theorem 13, Parts 3, 4 we get (note that F with $\delta(F) = 0$ is balanced linearly lean iff $M(F)$ is non-singular):

Lemma 17. *Consider a clause-set F with* $\delta(F) = 0$*. If F is balanced linearly lean, then F is balanced linear-autarky-decomposable if and only if F is a generalised sum of clause-sets* F_1, F_2 *with* $\delta(F_1) = \delta(F_2) = 0$*. Thus it is decidable in poly-time whether F is balanced linear-autarky-indecomposable.*

The *maximal deficiency* (see [5]) is defined as $\delta^*(F) = \max_{F' \subseteq F} \delta(F) \geq 0$ (note that $\delta(\top) = 0$). Analogously we define:

Definition 18. *The* ***maximal reduced deficiency*** *is defined as* $\boldsymbol{\delta^*_{\mathrm{r}}(F)} := \max_{F' \subseteq F} \delta_{\mathrm{r}}(F') \geq 0$.

If F is complement-invariant and F_0 is a core half of F, then $\delta^*_{\mathrm{r}}(F) = \delta^*(F_0)$, and thus for complement-invariant clause-sets $\delta^*_{\mathrm{r}}(F)$ is computable in polynomial time. By Lemma 7.2 in [6] we have:

Lemma 19. *If* F_0 *is balanced linearly lean, then* $\delta^*(F_0) = \delta(F_0)$ *(and thus* $\delta(F) \geq 0$*). It follows that for a linearly lean complement-invariant clause-set F we have* $\delta^*_{\mathrm{r}}(F) = \delta_{\mathrm{r}}(F)$ *(and thus* $\delta_{\mathrm{r}}(F) \geq 0$*).*

Theorem 20. *Assume Conjecture 8. For complement-invariant clause-sets F with* $\delta^*_{\mathrm{r}}(F) = 0$ *the lean kernel is computable in polynomial time (together with a maximal autarky realising the lean kernel). Thus the satisfiability problem is decidable in polynomial time (providing also a satisfying assignment in the satisfiable case), and furthermore if F is unsatisfiable, then a minimally unsatisfiable sub-clause-set can be computed in polynomial time.*

Proof. First the input F is reduced to the linearly lean kernel $F' \subseteq F$, where we have $\delta_{\mathrm{r}}(F') = 0$. If F' is lean, then F' is the lean kernel of F; otherwise by Conjecture 8 (and Lemma 9, Part 1) we can find a non-trivial autarky φ for F', and reduce F' to $F'' \subseteq F$. The whole cycle is repeated until we find the lean kernel of F. □

5 Hypergraph colouring

Considering hypergraphs G (that is, pairs $G = (V(G), E(G))$, where $V(G)$ is a (finite) set of vertices, and $E(G)$ is the hyperedge set, a set of subsets of $V(G)$) as special clause-sets, we naturally transfer the notion of deficiency:

Definition 21. *For a hypergraph G let the **deficiency** be defined as $\boldsymbol{\delta(G)} := |E(G)| - |V(G)|$, while the **maximal deficiency** is $\boldsymbol{\delta^*(G)} := \max_{G' \subseteq G} \delta(G)$, where by "$G' \subseteq G$" we denote sub-hypergraphs of G, that is, hypergraphs G' with $V(G') \subseteq V(G)$ and $E(G') \subseteq E(G)$.*

The maximal deficiency of a hypergraph is the maximum size of a matching in the bipartite hyperedge-vertex graph, and thus is computable in polynomial time. Core halves of complement-invariant PN-clause-sets F (where a PN-clause-set contains only positive and negative clauses, and no mixed cases) can be naturally considered as hypergraphs $G(F)$, which actually can be defined for arbitrary clause-sets as the **variable hypergraph**, that is $V(G(F)) = \mathrm{var}(F)$, while $E(G(F)) = \{\mathrm{var}(C) : C \in F\}$. It is easy to see that a complement-invariant PN-clause-set F is satisfiable resp. minimally unsatisfiable if and only if $G(F)$ is 2-colourable resp. minimally non-2-colourable (see Lemma 8.1 in [6] for the general statement, regarding k-colouring of hypergraphs), and thus from Theorem 20 we directly obtain:

Corollary 22. *Given Conjecture 8, the 2-colouring problem for hypergraphs G with $\delta^*(G) = 0$ can be solved in polynomial time.*

A very informative classification of minimally non-2-colourable *intersecting* square hypergraphs has been given in [11], allowing to replace the complicated algorithm underlying Corollary 22 by some form of simple pattern matching; see [6] for more details (and for the interpretation for minimally unsatisfiable complement-invariant clause-sets of reduced deficiency zero with the property, that every two different clauses have some variable in common).

6 Open problems

The main open problem (the case $k = 0$ is Theorem 20):

Conjecture 23. For fixed $k \in \mathbb{N}_0$ the satisfiability problem for complement-invariant clause-sets F with $\delta_{\mathrm{r}}^*(F) \leq k$ is decidable in polynomial time (and in the satisfiable cases also a satisfying assignment can be computed).

One can further ask here whether we have fixed-parameter tractability in k here. Conjecture 23 implies that for fixed k the hypergraph 2-colouring problem is decidable in polynomial time for hypergraphs G with $\delta^*(G) \leq k$. What about m-colourability for arbitrary $m \geq 2$?

What is the resolution complexity of minimally unsatisfiable complement-invariant clause-sets F with $\delta_{\mathrm{r}}(F) = 0$ ($\delta_{\mathrm{r}}(F) = k$) ?

References

1. Richard A. Brualdi and Herbert J. Ryser. *Combinatorial Matrix Theory*, volume 39 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, 1991. ISBN 0-521-32265-0; QA188.B78 1991.
2. Richard A. Brualdi and Bryan L. Shader. *Matrices of sign-solvable linear systems*, volume 116 of *Cambridge Tracts in Mathematics*. Cambridge University Press, 1995. ISBN 0-521-48296-8; QA188.B79.
3. Oliver Kullmann. Investigations on autark assignments. *Discrete Applied Mathematics*, 107:99–137, 2000.
4. Oliver Kullmann. On the use of autarkies for satisfiability decision. In Henry Kautz and Bart Selman, editors, *LICS 2001 Workshop on Theory and Applications of Satisfiability Testing (SAT 2001)*, volume 9 of *Electronic Notes in Discrete Mathematics (ENDM)*. Elsevier Science, June 2001.
5. Oliver Kullmann. Lean clause-sets: Generalizations of minimally unsatisfiable clause-sets. *Discrete Applied Mathematics*, 130:209–249, 2003.
6. Oliver Kullmann. Constraint satisfaction problems in clausal form: Autarkies, minimal unsatisfiability, and applications to hypergraph inequalities. In Nadia Creignou, Phokion Kolaitis, and Heribert Vollmer, editors, *Complexity of Constraints*, number 06401 in Dagstuhl Seminar Proceedings. Internationales Begegnungs- und Forschungszentrum fuer Informatik (IBFI), Schloss Dagstuhl, Germany, 2006. `http://drops.dagstuhl.de/opus/volltexte/2006/803`.
7. Oliver Kullmann, Inês Lynce, and João Marques-Silva. Categorisation of clauses in conjunctive normal forms: Minimally unsatisfiable sub-clause-sets and the lean kernel. In Armin Biere and Carla P. Gomes, editors, *Theory and Applications of Satisfiability Testing - SAT 2006*, volume 4121 of *Lecture Notes in Computer Science*, pages 22–35. Springer, 2006. ISBN 3-540-37206-7.
8. Oliver Kullmann, Victor W. Marek, and Mirosław Truszczyński. Computing autarkies and properties of the autarky monoid. In preparation, January 2007.
9. Christos H. Papadimitriou and David Wolfe. The complexity of facets resolved. *Journal of Computer and System Sciences*, 37:2–13, 1988.
10. Neil Robertson, Paul D. Seymour, and Robin Thomas. Permanents, Pfaffian orientations, and even directed circuits. *Annals of Mathematics*, 150:929–975, 1999.
11. Paul D. Seymour. On the two-colouring of hypergraphs. *The Quarterly Journal of Mathematics (Oxford University Press)*, 25:303–312, 1974.