

On NP-hardness of the clique partition : independence number gap recognition and related problems

Busygin, Stanislav.; Pasechnik, Dmitrii V.

2006

Busygina, S., & Pasechnik, D. V. (2006). On NP-hardness of the clique partition: independence number gap recognition and related problems. *Discrete Mathematics*, 306(4), 460–463.

<https://hdl.handle.net/10356/80117>

<https://doi.org/10.1016/j.disc.2006.01.004>

© 2006 Elsevier. This is the author created version of a work that has been peer reviewed and accepted for publication by *Discrete Mathematics*, Elsevier. It incorporates referee's comments but changes resulting from the publishing process, such as copyediting, structural formatting, may not be reflected in this document. The published version is available at: DOI [<http://dx.doi.org/10.1016/j.disc.2006.01.004>].

Downloaded on 21 Jul 2024 22:16:50 SGT

On NP -hardness of the clique partition—Independence number gap recognition and related problems

Stanislav Busygin^{a,*}, Dmitrii V. Pasechnik^{b,1}

^a*Industrial and Systems Engineering Department, University of Florida,
303 Weil Hall, Gainesville, FL 32611, USA*

^b*Theoretische Informatik, FB15, University of Frankfurt, Robert-Mayer Str. 11-15,
Postfach 11 19 32, 60054 Frankfurt am Main, Germany*

* Corresponding author. *E-mail addresses:* busygin@ufl.edu (S. Busygin),
dima@thi.informatik.uni-frankfurt.de (D.V. Pasechnik).

¹ Supported by the DFG Grant SCHN-503/2-1.

Abstract

We show that for a graph G it is NP -hard to decide whether its independence number $\alpha(G)$ equals its clique partition number $\overline{\chi}(G)$ even when some minimum clique partition of G is given. This implies that any $\alpha(G)$ -upper bound provably better than $\overline{\chi}(G)$ is NP -hard to compute.

To establish this result we use a reduction of the quasigroup completion problem (QCP, known to be NP -complete) to the maximum independent set problem. A QCP instance is satisfiable if and only if the independence number $\alpha(G)$ of the graph obtained within the reduction is equal to the number of holes h in the QCP instance. At the same time, the inequality $\overline{\chi}(G) \leq h$ always holds. Thus, QCP is satisfiable if and only if $\alpha(G) = \overline{\chi}(G) = h$. Computing the Lovász number $\vartheta(G)$ we can detect QCP unsatisfiability at least when $\overline{\chi}(G) < h$. In the other cases QCP reduces to $\overline{\chi}(G) - \alpha(G) > 0$ gap recognition, with one minimum clique partition of G known.

Keywords

Independence number; Clique partition number; Lovász number; Latin square; Quasigroup completion problem

1. Introduction

Let $G(V, E)$ be a simple undirected graph. An *independent set* of vertices is a subset $S \subseteq V$ such that any two vertices of S are *not* adjacent. The *maximum independent set problem* asks for an independent set of the maximum cardinality. This cardinality $\alpha(G)$ is called the *independence number* of the graph, and is NP -hard to compute [5]. A *clique* Q is a subset of V such that any two vertices of Q are adjacent. The *minimum clique partition problem* asks for a smallest by cardinality set of cliques $\{Q_1, \dots, Q_{\overline{\chi}}\}$ containing every vertex $v \in V$ in exactly one of the cliques. The cardinality $\overline{\chi}(G)$ of this set is called the *clique partition number*. It is equal to the *chromatic number* $X(\overline{G})$ (minimum number of vertex colors needed to provide different colors for any pair of adjacent vertices) of the complementary graph. The minimum clique partition problem is also NP -hard [5].

Obviously, the inequality $\alpha(G) \leq \bar{\chi}(G)$ holds as no two vertices of an independent set can belong to the same clique.

There exists a polynomial-time computable function $\vartheta(G)$ “sandwiched” between those two *NP*-hard numbers [13,12]:

$$\alpha(G) \leq \vartheta(G) \leq \bar{\chi}(G). \quad (1)$$

One simple definition of $\vartheta(G)$ is via minimum of the largest eigenvalue of so-called *feasible matrices* $A = (a_{ij})_{n \times n}$:

$$\vartheta(G) = \min_A \lambda_{\max}(A), \quad (2)$$

$$\text{s.t. } a_{ij} = 1 \quad \text{if } (i, j) \notin E; \quad a_{ij} = a_{ji},$$

(that is, to obtain $\vartheta(G)$ we minimize the largest eigenvalue of a symmetric matrix having 1s on the main diagonal and in all entries corresponding to non-edges, while the other entries are arbitrary). $\vartheta(G)$ is called the *Lovász number* (or *ϑ -function*) of a graph. It serves as an upper bound for the independence number and as a lower bound for the clique partition number simultaneously. For the numerous application of the Lovász number in optimization see, e.g., [6]. Besides, there are increasingly tight sequences of polynomial-time computable upper bounds for $\alpha(G)$ based on “lift-and-project” method [14] and the concept of matrix copositivity [4].

A *latin square* is an $n \times n$ matrix filled with integers from 1 to n so that each number occurs exactly once in any row and in any column. One example is $L = (\ell_{ij})_{n \times n}$ such that

$$\ell_{ij} = ((i + j - 2) \bmod n) + 1. \quad (3)$$

In the *quasigroup completion problem* (QCP, a.k.a. latin square completion) we are given an $n \times n$ array partially filled with integers from $\{1, \dots, n\}$ and it is asked whether there is a completion for all h empty cells (*holes*) such that it gives a latin square. QCP is *NP*-complete [3]. Recently it has been intensively studied, especially from constraint programming and boolean satisfiability viewpoints [8,15,1,11,9,10,7].

In this paper, we show a reduction of QCP to the maximum independent set problem. The obtained graph instances obey $\alpha(G) \leq h$ and $\bar{\chi}(G) \leq h$ constraints. At that, the original QCP instance is satisfiable if and only if $\alpha(G) = \bar{\chi}(G) = h$. This allows us to obtain some results restricting polynomial-time recognition of $\bar{\chi}(G) - \alpha(G) > 0$ gap and computation of tight upper bounds on $\alpha(G)$ under $P \neq NP$ assumption. In fact, unless $P = NP$, it means there is no polynomial-time computable upper bound on $\alpha(G)$ provably better than $\vartheta(G)$ and, in turn, $\vartheta(G)$ does not bound $\alpha(G)$ provably better than $\bar{\chi}(G)$.

Clearly, all the presented complexity results also hold for the chromatic-clique number gap recognition by taking the complementary graph.

2. Latin square 3D encoding

The concept of latin square can be also expressed via a three-dimensional array of 0-1 values. Namely, let a 0-1 variable x_{ijk} , $i, j, k \in \{1, \dots, n\}$, denote “Cell (i, j) is filled with number k ”. The array of these variables determines a latin square if and only if

$$\begin{cases} \forall i, j & \sum_{k=1}^n x_{ijk} = 1, \\ \forall i, k & \sum_{j=1}^n x_{ijk} = 1, \\ \forall j, k & \sum_{i=1}^n x_{ijk} = 1. \end{cases} \quad (4)$$

These conditions correspond to maximum independent sets of a graph, whose vertices are triples (i, j, k) and there is an edge between two of them if and only if two of their entries coincide. This graph Γ is known as $H(3, n)$ Hamming graph, see e.g. [2].

Lemma 1. $\alpha(\Gamma) = n^2$. There is a one-to-one correspondence between maximum independent sets of Γ and $n \times n$ latin squares.

Proof. First, we prove Γ does not have an independent set larger in size than n^2 . Indeed, there are only n^2 distinct pairs of two first entries (i, j) for the vertices $\{(i, j, k)\}$. Thus, in any vertices subset X such that $|X| > n^2$ there is at least one vertex pair $((i, j, k), (p, q, r))$ such that $(i, j) = (p, q)$. This vertex pair must be connected by an edge. So, X is not an independent set.

Now, consider a latin square $L = (\ell_{ij})_{n \times n}$ and the vertex subset $S = \{(i, j, k) : \ell_{ij} = k\}$. It contains n^2 vertices because there are n^2 distinct (i, j) pairs. Let $(i, j, k) \in S$ and $(p, q, r) \in S$ be two distinct vertices. As $i = p$ and $j = q$ would have implied $k = r$ by the definition of S , this case is not possible. Thus, if $i = p$, then $j \neq q$ and $k \neq r$ as L does not have two equal numbers on the same row. Similarly, if $j = q$, then $i \neq p$ and $k \neq r$ as L does not have two equal numbers on the same column. Therefore, there are no triples in S with exactly two common entries and hence L defines a maximum independent set of Γ . As (3) provides a latin square for any $n > 0$, one obtains $\alpha(\Gamma) = n^2$ as claimed.

Conversely, it is easy to see that any maximum independent set $S = \{(i, j, k)\}$, $|S| = n^2$ defines a latin square $L = (\ell_{ij})_{n \times n}$ such that $\ell_{ij} = k$ if and only if $(i, j, k) \in S$. \square

To reduce QCP to the maximum independent set problem we will use subgraphs of Γ . Let the QCP input be a matrix $L = (\ell_{ij})_{n \times n}$ such that $\ell_{ij} = k \in \{1, \dots, n\}$ if the cell (i, j) is prefilled with k , and $\ell_{ij} = 0$ otherwise. Correspondingly, the number of holes h is the total number of entries (i, j) such that $\ell_{ij} = 0$. Without loss of generality we assume that this input does not immediately violate the latin square constraints. That is, $\ell_{ij} = \ell_{iq} > 0$, $j \neq q$, or $\ell_{ij} = \ell_{pj} > 0$, $i \neq p$, cases never occur. Otherwise, the QCP instance is trivially unsatisfiable. Define a graph $G(V, E)$ with vertices

$$V = \{(i, j, k) : (\ell_{ij} = 0) \& (\forall p : \ell_{pj} \neq k) \& (\forall q : \ell_{iq} \neq k)\}. \quad (5)$$

As earlier, put an edge between distinct vertices (i, j, k) and (p, q, r) when they have two common entries:

$$E = \{((i, j, k), (p, q, r)) : (i = p) \& (j = q) \& (k \neq r) \vee (i = p) \& (j \neq q) \& (k = r) \vee (i \neq p) \& (j = q) \& (k = r)\}. \quad (6)$$

In other words, $G(V, E)$ is the subgraph of Γ induced by non-neighbors of those vertices (i, j, k) for which $\ell_{ij} = k > 0$.

Lemma 2. $\alpha(G) \leq h$. *The QCP instance given by the matrix L is satisfiable if and only if $\alpha(G) = h$.*

Proof. Let $S_0 = \{(i, j, k) : \ell_{ij} = k > 0\}$ be the vertex subset of Γ corresponding to the partial completion given by L . Obviously, $|S_0| = n^2 - h$. Since the partial completion obeys the latin square constraints, S_0 is an independent set. Denote by $N^+(S_0)$ the closed neighborhood of S_0 , that is, union of S_0 with the set of vertices of Γ adjacent to at least one vertex from S_0 . $G(V, E)$ is obtained by removing $N^+(S_0)$ from Γ .

Assume $G(V, E)$ has an independent set S_1 of size greater than h . Then $S_0 \cup S_1$ is an independent set of Γ having more than n^2 vertices, contradicting Lemma 1. Hence $\alpha(G) \leq h$.

Let $G(V, E)$ have a maximum independent set S_1 of size h . Then $S_0 \cup S_1$ is a maximum independent set of Γ , so S_1 determines a correct completion of the QCP input to a latin square. Therefore, the given QCP instance is satisfiable. Conversely, if the given input matrix L admits a completion to a latin square, we can take the maximum independent set S of Γ corresponding to this latin square and observe that $S \setminus S_0$ is an independent set of $G(V, E)$ of size h . Therefore, the QCP instance is satisfiable if and only if $\alpha(G) = h$. \square

Thus, we have described a reduction of QCP to the maximum independent set problem on subgraphs of Γ . The next section concerns their clique partitions.

3. The main results

Lemma 3. *Let $G(V, E)$ be a graph obtained within the QCP reduction to the maximum independent set problem. Then $\bar{\chi}(G) \leq h$.*

Proof. Let $L = (\ell_{ij})_{n \times n}$ be the QCP input matrix as described above. V may include only such vertices (i, j, k) for which $\ell_{ij} = 0$. We note that all $(i, j, k) \in V$ corresponding to one hole $\ell_{ij} = 0$ comprise a clique. Hence, V is a union of not more than h of such cliques. This implies $\bar{\chi}(G) \leq h$. \square

Therefore, computing the Lovász number $\vartheta(G)$ on the described graphs we can efficiently detect QCP unsatisfiability at least when $\bar{\chi}(G) < h$. We may say that the inequality $\vartheta(G) < h - \varepsilon$ for some fixed $0 < \varepsilon < 1$ designates an easily recognizable

subclass of unsatisfiable QCP instances. In the other cases, QCP is equivalent to deciding whether $\alpha(G) = \bar{\chi}(G)$ provided $\bar{\chi}(G) = h$ and the clique partition defined in the proof of Lemma 3 is a minimum one. Thus, we have deduced the following:

Theorem 4. *For a graph G is it NP-hard to decide whether there is a gap between its independence and clique partition numbers $\bar{\chi}(G) - \alpha(G) > 0$ provided some minimum clique partition of G is given.*

We note that currently we are not aware of any graph G obtained within the reduction from an unsatisfiable QCP instance for which $\vartheta(G) = \bar{\chi}(G) = h$.

Corollary 5. *For a graph G is it NP-hard to decide whether there is a gap between its independence and clique partition numbers $(G) - \alpha(G) > 0$.*

Though it immediately follows from Theorem 4, there is also a simple direct proof of this fact. Assume we have an oracle answering whether $\bar{\chi}(G) - \alpha(G) > 0$ for any graph G . Define G_i as the graph composed of G and i additional mutually independent vertices, each of which is connected with every vertex of G . Note that $\alpha(G_i) = \max(\alpha(G), i)$ and $\bar{\chi}(G_i) = \max(\bar{\chi}(G), i)$. Submit the graphs G_i , $i = 0, 1, \dots$, to the oracle until it says $\alpha(G_i) = \bar{\chi}(G_i)$. There cannot be more than $\bar{\chi}(G)$ of such queries. Upon termination, $\vartheta(G_i) = \bar{\chi}(G)$ since $\alpha(G_i) = \bar{\chi}(G_i) = \bar{\chi}(G)$, so using the oracle we can compute $\bar{\chi}(G)$ in polynomial time. (In fact, we have to compute $\vartheta(G_i)$ only if the process stops with $i = 0$, that is, when $\alpha(G) = \bar{\chi}(G)$. Otherwise the terminal value i gives $\bar{\chi}(G)$.)

Corollary 6. *Unless $P = NP$, there is no polynomial-time computable upper bound on the independence number $\alpha(G)$ provably better than the Lovász number $\vartheta(G)$ and, in turn, $\vartheta(G)$ bounds $\alpha(G)$ from above not provably better than the clique partition number $\bar{\chi}(G)$.*

Indeed, any such upper bound on $\alpha(G)$ allows for polynomial time recognition of $\bar{\chi}(G) - \alpha(G) > 0$ gap whenever $\bar{\chi}(G)$ is known. According to Theorem 4, this would imply $P = NP$.

References

- [1] D. Achlioptas, C. Gomes, H. Kautz, B. Selman, Generating satisfiable problem instances, in: Proceedings of the 17th National Conference on Artificial Intelligence (AAAI-00), Austin, TX, 2000.
- [2] E. Bannai, T. Ito, Algebraic Combinatorics I: Association Schemes, Benjamin/Cummings, Menlo Park, CA, 1984.
- [3] C. Colbourn, The complexity of completing partial latin squares, *Discrete Appl. Math.* 8 (1984) 25–30.
- [4] E. de Klerk, D. Pasechnik, Approximation of the stability number of a graph via copositive programming, *SIAM J. Optim.* 12 (4) (2001) 875–892.
- [5] M. Garey, D. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, Freeman & Co., New York, 1979.
- [6] M.X. Goemans, Semidefinite programming in combinatorial optimization, *Math. Programming* 79 (1997).
- [7] C. Gomes, R. Regis, D. Shmoys, An improved approximation for the partial latin square extension problem, in: Proceedings of the ACM-SIAM Symposium on Discrete Algorithms (SODA-2003), Baltimore, MD, 2003.
- [8] C. Gomes, B. Selman, Problem structure in the presence of perturbations, in: Proceedings of the 14th National Conference on Artificial Intelligence (AAAI-97), New Providence, RI, 1997.
- [9] C. Gomes, D. Shmoys, Completing quasigroups or latin squares: a structured graph coloring problem, in: Proceedings of the Computational Symposium on Graph Coloring and Extensions, 2002.
- [10] C. Gomes, D. Shmoys, The promise of LP to boost CSP techniques for combinatorial problems, in: Proceedings of the 4th International Symposium on Integration of AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems (CP-AI-OR'02), Le Croisic, France, 2002, pp. 291–305.
- [11] H. Kautz, Y. Ruan, D. Achlioptas, C. Gomes, B. Selman, M. Stickel, Balance and filtering in structured satisfiable problems, in: Proceedings of the 17th International Conference on Artificial Intelligence (IJCAI-2001), Seattle, WA, 2001.
- [12] D.E. Knuth, The Sandwich theorem, *Electron. J. Combin.* 1 (1994).
- [13] L. Lovász, On the Shannon capacity of a graph, *IEEE Trans. Inform. Theory* 25 (1) (1979) 1–7.
- [14] L. Lovász, A. Schrijver, Cones of matrices and set-functions and 0-1 optimization, *SIAM J. Optim.* 1 (2) (1991) 166–190.
- [15] I. Peterson, Completing latin squares, *Sci. News* 19 (157) (2000).