A lower bound for approximating the grundy number

Guy Kortsarz[1†]

received 1, revised 14th December 2006, accepted 1.

The grundy number of a graph is the maximum number of colors used by on-line first-fit coloring, under the worst order of arrival of vertices. The grundy number problem is to find this ordering. We prove that there is a constant $c > 1$ so that approximating the grundy number problem within c is not possible, unless $NP \subseteq RP$

Keywords: Grundy, number, coloring, approximation, hardness

[†]Rutgers University, Camden

Contents

1 Introduction

In first-fit coloring, we color the vertices with positive integers. The vertices arrive on-line, and we color the next vertex v by the minimum number not appearing in its neighborhood. The grundy number of a graph is the maximum possible number of colors used by on-line first-fit coloring, under the worst order of arrival of vertices. The grundy number problem is to find this worst ordering. The problem was extensively studied, see e.g., **??????**. The first-fit coloring is important in practice and is extensively studied as well, see, e.g., **??**. In **?** the author studies the grundy number problem on interval graphs and establishes that the grundy number differs from the obvious lower bound of maximum clique size by at most some multiplicative universal constant. An algorithm solving this problem on trees is given in **?** and on partial k-trees in **?**. The problem was shown to be NP-complete on directed graphs **?**. The NP-hardness for the undirected case was harder to prove. Indeed, in the book **?** proving NP-hardness for the grundy number problem is posed as an open problem. A proof of the NP-hardness of the undirected version was established slightly after that **?**. Prior to this paper, there was no hardness of approximation result for the problem.

1.1 Definitions

An *independent set* in $G(V, E)$ is a subset $V' \subseteq V$ so that no two vertices in V' are neighbors. A dominating set is a subset $V' \subseteq V$ so that $V' \cup N(V') = V$ with $N(V')$ the vertices not in V' with at least one neighbor in V'. An independent set is (inclusion) maximal if it is also a dominating set of the graph. A set U dominates Z if every vertex $z \in Z$ either belongs to U or has at least one neighbor in U. A grundy coloring is a disjoint partition $V = C_1 \cup C_2 \ldots \cup C_k$ of V into independent sets so that every $v \in C_j$ has at least one neighbor in every C_i, $i < j$. Thus, C_i should be a maximal independent set in the graph induced by $\bigcup_{j \geq i} C_j$. The C_i sets are also called the *color-classes* of the partition. The goal is to maximize k, the number of color-classes. The largest possible k is the grundy number of the graph.

1.2 Related problems

A natural idea for approximating the grundy number problem is trying to find a small size maximal independent set, removing the vertices of this independent set, and iterating. Unfortunately, in **?** it is shown that for every $\epsilon > 0$ approximating the size of the minimum size maximal independent set within $n^{1-\epsilon}$ is $NP-$hard. In fact, the author is not aware of any non-trivial approximation published for the grundy number problem. Conversely, one may hope that **?** could be used to derive a strong lower bound. However, if a graph has a large grundy number, it does not necessarily imply that all maximal independent sets are small. For example, it might be the case that $|C_1| = n/2$ and, $|C_j| = 2$ for every $j \geq 2$ and so the grundy number equals $\Omega(n)$ while there exists a a maximal independent set of size $\Omega(n)$. The grundy number problem is thus more related to the following problem: given a size bound x, find an independent set I of size at most x, and maximize $|N(I)|$. It is possible to show that this problem cannot be approximated within some constant $c > 1$, unless $P = NP$ (detail omitted). Still, it seems hard to deduce from that any hardness result for the grundy number problem. This is as a large grundy number requires a large collection of small independent sets with many neighbors each.

Other two maximization partitioning problems that have some relation to the grundy number problem are the achromatic number problem and the complete coloring problem (cf. **?** for a survey). The achromatic number problem is to partition the vertex set of a graph $G(V, E)$ to maximum number k of independent sets $V = \bigcup_{i=1}^{k} C_i$ so that every two independent sets in the partition share an edge (for every two independent sets C_i, C_j in the partition there exist a vertex $v \in C_i$ and a vertex $u \in C_j$ so that $(u, v) \in E$). The complete partition problem is defined in the same way as the achromatic number problem except that the C_i sets need not be independent sets and can be any arbitrary subsets of V.

Even though the grundy number problem seems more important in practice than the achromatic and complete partition problems (because of its tight relation to first-fit coloring), an $\Omega(\sqrt{\log n})$ lower bound on the approximability of the two latter problems is known **?**, but no lower bound for approximating the grundy number problem was known before this paper. It may be that grundy number problem is harder to analyze than the two other problems, because of its more restricted nature: Every vertex in C_j $j > i$ has to be dominated by C_i (in the other two problems, C_i, C_j need to share only one edge). This poses a difficulty in the reduction, in particular for the part of the reduction of a yes-instance. The construction of **?** does not seem to work. The construction here is different. The technical nature of the proofs is also different.

1.3 Background

We now describe a construction from **?**. See also **???** for constructions using similar tools.

Let $\mathcal{I}_n$ be the set of 5-regular, simple graphs $I = (\mathcal{V}(I), \mathcal{E}(I))$ with $\mathcal{V} = \{1, 2, \ldots, n\}$. We say that a graph $I \in \mathcal{I}_n$ is ϵ-non-3-colorable if in every 3-coloring of the vertices of I at most $(1 - \epsilon)|\mathcal{E}(I)|$ edges of I are properly colored.

The maximum legal edge 3-coloring problem is given some graph G, to legally color the vertices of V and to maximize the number of edges that are legally colored. The following theorem states that given any NP-complete instance, then there exists a universal constant ϵ and a polynomial time reduction to an instance of the maximum legal edge 3-coloring problem so that in a yes instance, the graph is 3-colorable and in a no instance at most $1 - \epsilon$ fraction of the edges are 3-colorable.

Theorem 1.1 **?** *There is an $\epsilon > 0$ such that the following holds. Let $L \subseteq \{0, 1\}^*$ be a language in* NP. *Then, there is a polynomial computable function $f : \{0, 1\}^* \to \mathcal{I}_n$ such that for all $x \in \{0, 1\}^*$*

- *if $x \in L$, then $f(x)$ is 3-colorable;*
- *if $x \notin L$, then $f(x)$ is ϵ-non-3-colorable.*

The proof of the theorem in **?** is based on an earlier reduction of Petrank from an arbitrary NP language to an instance of 3-SAT (see **?**). The advantage of reducing to maximum legal edge coloring is that by permuting the colors, if one legal coloring exists, many *distinct* colorings can be derived.

Notation 1.2 *Throughout the paper,* $n = |\mathcal{V}|$ *denotes the number of vertices in* I.

Consider the following two-prover protocol.

Input: $I \in \mathcal{I}_n$.

Verifier: Chooses a $v \in \mathcal{V}(I)$ at random. Chooses a random pair of edges (e, e'), both incident on v. Sends e to the Prover I and e' to Prover II.

Prover I: Sends a proper 3-coloring $\chi(e)$ to the verifier.

Prover II: Sends a proper 3-coloring $\chi'(e')$ to the verifier.

Verifier: Accepts iff $\chi(v) = \chi'(v)$.

Proposition 1.3 **?** *Let the maximum number of edges properly colored in any 3-coloring of* I *be* $(1 - \epsilon)|\mathcal{E}(I)|$. *Let* Π *be a strategy for the provers that makes the verifier accept with the highest probability. Then, under* Π, $\Pr[$*Verifier accepts*$] = 1 - \epsilon$.

We briefly explain why does the proposition hold. In case of a yes instance, the two provers may use the same legal 3-coloring deriving a probability 1 of acceptance by the verifier. Now, given a no instance, every prover has some fix strategy (namely choice of colors) over the edges. This is as the provers are deterministic. By Theorem 1.1, the colorings of the two provers differ on an ϵ fraction of the edges. Since the graph is regular, it follows from the method used that the selected edge is random as well. Hence with probability at least ϵ, the random edge e selected is one of those edges that have two different colorings under the strategy of the provers. For such an edge e, the verifier rejects. Hence, the verifier rejects with probability at least ϵ.

Let A_e denote some question to the first prover (in our case a question is simply the edge e to be colored) and $B_{e'}$ a question to the second prover. If there is non-zero probability that $(A_e, B_{e'})$ are asked together, namely, if e, e' share a vertex, then we call $(A_e, B_{e'})$ a *query*.

Consider the following ℓ-fold parallel repetition of the protocol above.

Input: $I \in \mathcal{I}_n$.

Verifier: Sends $v_1, v_2, \ldots, v_\ell \in \mathcal{V}(I)$ at random. For each $i = 1, 2, \ldots, \ell$, chooses a random pair (e_i, e'_i) of edges both incident on v_i. Sends $\langle e_1, e_2, \ldots, e_\ell \rangle$ to the first prover and $\langle e'_1, e'_2, \ldots, e'_\ell \rangle$ to the second prover.

Prover I: Chooses $\langle \chi_1, \chi_2, \ldots, \chi_\ell \rangle$ to the verifier, where $\chi_i(e_i)$ is a proper coloring of e_i.

Prover II: Sends $\langle \chi'_1, \chi'_2, \ldots, \chi'_\ell \rangle$ to the verifier, where $\chi_i(e'_i)$ is a proper coloring of e'_i.

Verifier: Accepts iff for all $i = 1, 2, \ldots, \ell$, $\chi_i(v_i) = \chi'_i(v_i)$.

Theorem 1.4 (Raz's parallel repetition theorem ?) *If we perform ℓ parallel repetitions of a one-round two-prover protocol with probability of acceptance ϵ, and accept if and only if all ℓ runs lead to acceptance, then the probability of acceptance is at most $2^{-c\ell}$ where $c > 0$ is a constant that depends only on ϵ.*

Corollary 1.5 *Let the maximum number of properly colored edges in any 3-coloring of I be $(1-\epsilon)|\mathcal{E}(I)|$ ($\epsilon > 0$). Then, there exists a constant $c = c(\epsilon) > 0$ such that the following holds for all ℓ. Let Π be a strategy for the provers in the above protocol that makes the verifier accept with the highest probability. Then, under Π, $\Pr[\text{Verifier accepts}] \leq 2^{-c\ell}$.*

From now on, we will use c for the constant associated in the above corollary with the constant ϵ.

Remark: In this paper ℓ will equal some large enough universal constant.

A query for the parallel repetition protocol is defined as above; namely, it is two tuples of ℓ questions each that have positive probability of being sent together. The task of determining the best strategy for the provers in the parallel protocol is now posed in graph theoretic terms as follows. We construct a bipartite graph $\mathsf{MRG}(I) = (U_1, U_2, E)$. The set $U_1 = \bigcup A_{\overline{e}}$ is decomposed to "question clusters" $\{A_{\overline{e}}\}$. Intuitively, every $A_{\overline{e}}$ is a possible question (tuple of ℓ edges) for the first prover. The set $U_2 = \bigcup B_{\overline{e'}}$ is a union of sets $B_{\overline{e'}}$. Every $B_{\overline{e'}}$ corresponds to a question for the second prover. The actual vertices in $\mathsf{MRG}(I)$ are all the proper answers (namely, legal colorings) for every $A_{\overline{e}}$ and $B_{\overline{e'}}$. We call these vertices *answer-vertices*. If $(A_{\overline{e}}, B_{\overline{e'}})$ is not a query no $x \in A_{\overline{e}}$ and $y \in B_{\overline{e'}}$ share an edge in $MRG(I)$. If $(A_{\overline{e}}, B_{\overline{e'}})$ is a query, then answer-vertices $x \in A$ and $y \in B$ are adjacent in $\mathsf{MRG}(I)$ (or $(x, y) \in E$) iff they are consistent, namely assign the same color to the shared vertex in every edge in the $\ell-$tuple. Formally:

Definition 1.1 (min-rep-Graph ?) *Given a graph $I \in \mathcal{I}_n$, the bipartite graph $\mathsf{MRG}(I) = (U_1, U_2, E)$, is defined as follows. U_1 and U_2 consist of tuples*

$$\langle e_1, e_2, \ldots, e_\ell, \chi_1, \chi_2, \ldots, \chi_\ell \rangle,$$

where for $i = 1, 2, \ldots, \ell$, $e_i \in \mathcal{E}(I)$ and $\chi_i(e_i)$ is a proper coloring of e_i. The edges of $\mathsf{MRG}(I)$ will be labeled by elements of $\mathcal{V}^\ell$, and we will use $(u, u', \overline{v})$ to denote the edge between $u \in U_1$ and $u' \in U_2$ with label $\overline{v} \in \mathcal{V}(I)^\ell$. Let $E_{\overline{v}}$ be the edges labeled $\overline{v}$.

$$\begin{aligned} E_{\overline{v}} \quad = \quad & \{(\langle \overline{e}, \overline{\chi} \rangle, \langle \overline{e'}, \overline{\chi'} \rangle, \overline{v}) : \langle \overline{e}, \overline{\chi} \rangle \in U_1 \text{ and } \langle \overline{e'}, \overline{\chi'} \rangle \in U_2 \text{ and} \\ & \text{for } i = 1, 2, \ldots, \ell,\ v_i \in e_i, e'_i \text{ and } \chi_i(v_i) = \chi'_i(v_i)\}. \end{aligned}$$

Finally, the edge set of $\mathsf{MRG}(I)$ is $\bigcup_{\overline{v} \in \mathcal{V}(I)^\ell} E_{\overline{v}}$.

Note: In the graph above, and in those we shall encounter below, we allow parallel edges, but different edges connecting the same two vertices have different labels.

The vertices of U_1 can be partitioned into sets $(A_{\overline{e}} : \overline{e} \in \mathcal{E}(I)^\ell)$, where $A_{\overline{e}}$ consists of vertices of the form $\langle \overline{e}, \overline{\chi} \rangle$. Similarly, U_2 can be partitioned into sets $(B_{\overline{e'}} : \overline{e'} \in \mathcal{E}(I)^\ell)$. Note that $|A_{\overline{e}}| = |B_{\overline{e'}}| = 6^\ell$, and there are precisely $q = |\mathcal{E}(I)^\ell| = (\frac{5}{2})^\ell |\mathcal{V}(I)|^\ell$ parts in each partition (recall that the graph is 5-regular).

Definition 1.2 (Super-graph) *The supergraph $SG(I)$ of the protocol has vertex set $\{A_{\overline{e}}\}$ and $\{B_{\overline{e'}}\}$. A pair $\tilde{f} = (A_{\overline{e}}, B_{\overline{e'}})$ is a superedge iff it is a query (namely, for $i = 1, 2, \ldots, \ell$, e_i and e'_i have a common vertex).*

Remark: To avoid confusion, we use $\tilde{f}$ for superedges and $\overline{e}$ for edge-tuples of size ℓ of $\mathcal{E}(I)$.

The number of superedges is denoted by h. The supergraph is d-regular with $d = 2^\ell \cdot 5^\ell = 10^\ell$; indeed, consider a supervertex $A_{\overline{e}}$ and count its neighbors in $SG(I)$. A neighbor $B_{\overline{e'}}$ of $A_{\overline{e}}$ has the same $\bar{v}$ tuple. The tuple $\overline{v}$ can be chosen in 2^ℓ ways. Once this tuple is chosen, for defining $B_{\overline{e}}$, in every coordinate, there are 5 ways to choose edges touching the respective vertex in $\bar{v}$ vertex (this follows as the graph is 5-regular). This implies that $d = 10^\ell$. Also, $h = q \cdot d = (25 \cdot n)^\ell$.

Definition 1.3 *Let $S \subseteq U_1 \cup U_2$. We say that a superedge $\tilde{f} = (A_{\overline{e}}, B_{\overline{e'}})$ is* min-rep covered *by S if there exist $a \in A_{\overline{e}} \cap S$ and $b \in B_{\overline{e'}} \cap S$ that are adjacent in $\mathsf{MRG}(I)$ (or $(a,b) \in E$).*

A set S is a min-rep cover if it covers all the superedges.

We are now ready to derive the key properties of $\mathsf{MRG}(I)$ as established by **?**.

Lemma 1.6 (Main Lemma) *We have the following properties depending on whether I is 3-colorable or whether I is ϵ-non-3-colorable.*

- *If $I \in \mathcal{I}_n$ is 3-colorable, then $V(\mathsf{MRG}(I))$ can be partitioned into 6^ℓ disjoint min-rep covers each of size $2q$. For every $\overline{e}$, $\overline{e'}$ each one of the min-rep covers contains exactly one vertex from every $A_{\overline{e}}$, $B_{\overline{e'}}$.*
- *If I is ϵ-non-3-colorable, then for any S of size $2qk$, at most an α-fraction of edges of $\mathsf{MRG}(I)$ are min-rep covered, where $\alpha \le (32k^2 2^{-c\ell})^{1/3}$.*

The graph induced by $A_{\overline{e}} \cup B_{\overline{e'}}$: Say that $\tilde{f} = (A_{\overline{e}}, B_{\overline{e'}})$ is a superedge (namely $(A_{\overline{e}}, B_{\overline{e'}})$ is a query). Consider the graph induced by $A_{\overline{e}} \cup B_{\overline{e'}}$ in $MRG(I)$. The set $A_{\overline{e}}$ has 6^ℓ vertices, each corresponding to assigning a different coloring to the edges involved in $\overline{e}$. Some of these colorings agree on what they assign to the vertices in $\overline{v}$. These vertices are put into a single class. This gives 3^ℓ classes of size 2^ℓ each. The size of any class is 2^ℓ because given the colors of $\overline{v}$, there are two proper ways to complete the coloring of the other vertices in the edges of $\overline{e}$ on every coordinate. Similarly, we partition $B_{\overline{e'}}$ into 3^ℓ classes of size 2^ℓ each. By the definition of $\mathsf{MRG}(I)$, two classes $C \subseteq A_{\overline{e}}$, $D \subseteq B_{\overline{e'}}$ that agree on the coloring for $\overline{v}$ form a complete bipartite graph. Otherwise, the graph induced by $C \cup D$ is an independent set. Thus, the graph induced by $A_{\overline{e}} \cup B_{\overline{e'}}$ is a union of vertex disjoint complete bipartite graphs. See Figure 1.3.

Definition 1.4 *Let $\tilde{f} = (A_{\overline{e}}, B_{\overline{e'}})$ be a superedge. Let $C \subset A_{\overline{e}}$ be a 2^ℓ size class of vertices of $A_{\overline{e}}$ that agree on the colors to $\overline{v}$. The size 2^ℓ class $\mathcal{H}_{\tilde{f}}(C) \subseteq B_{\overline{e'}}$ giving the same colors to $\overline{v}$ is called the* matching class *of C for the superedge $\tilde{f} = (A_{\overline{e}}, B_{\overline{e'}})$ and is denoted $\mathcal{H}_{\tilde{f}}(C)$.*

1.4 The construction

Like **?** we derive the reduction by modifying $\mathsf{MRG}(I)$. However, the modification is different.

The intermediate random graph $\tilde{G}(I)$: The intermediate graph $\tilde{G}(I)$ is a random bipartite graph with vertex sets $\tilde{V}_1$ and $\tilde{V}_2$ similar to that of **???**. The first vertex set $\tilde{V}_1 = U_1 \cup U_2$ will, in fact, be the same as $V(\mathsf{MRG}(I))$. We call the vertices of $\tilde{V}_1$ *answer-vertices*, and as before, we partition it into $2q$ clusters $(A_{\overline{e}}, B_{\overline{e'}} : \overline{e}, \overline{e'} \in \mathcal{E}(I)^\ell)$ each, with 6^ℓ answer-vertices. The other vertex set $\tilde{V}_2$ has one *ground set* $M_{\tilde{f}}$ of $m \ge n^{2\ell}$ vertices for each superedge $\tilde{f}$ of $SG(I)$. In addition, we assume that 6^ℓ divides m (we increase m to the first integer larger than $n^{2\cdot\ell}$ so that $6^\ell | m$). The set $\tilde{V}_2$ has $h \cdot m = (25 \cdot n)^\ell \cdot m$ vertices.

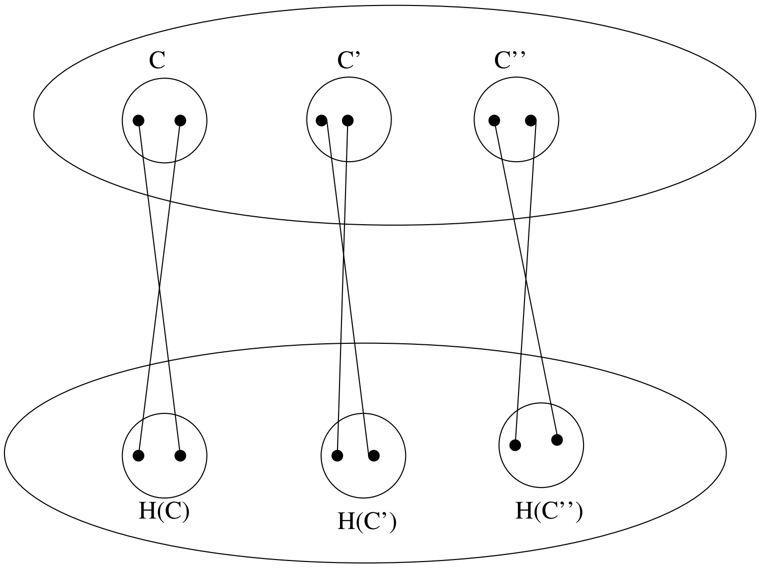

Fig. 1: The graph induced by $A_{\overline{e}} \cup B_{\overline{e'}}$ for $\ell = 1$. There are 6 vertices split into 3 classes. Each class has 2 vertices. This graph is a union of vertex-disjoint complete bipartite graphs.

We now describe how the edges are placed in $\tilde{G}(I)$. The edges connecting the ground set $M_{\tilde{f}}$ to the answer-vertices in $\tilde{V}_1$ are determined by using a random process. This process is run independently for each superedge $\tilde{f}$ of $SG(I)$. Consider a superedge $\tilde{f} = (A_{\overline{e}}, B_{\overline{e'}}, \overline{v})$ of $SG(I)$. The vertices in the ground set $M_{\tilde{f}}$ have neighbors only in $A_{\overline{e}}$ and $B_{\overline{e'}}$.

The neighbors in $M_{\tilde{f}}$ of a class $C \subseteq A_{\overline{e}}$ and its matching class $\mathcal{H}_{\tilde{f}}(C) \subseteq B_{\overline{e'}}$ are defined together and complement each other (see **?** for the first time this idea was used). Choose a random half $M_{\tilde{f}}(C)$ of $M_{\tilde{f}}$. Join all the answer-vertices of C to all the vertices of $M_{\tilde{f}}(C)$. Join every answer-vertex of $\mathcal{H}_{\tilde{f}}(C)$ to the complementary half $M_{\tilde{f}}(\bar{C}) = M_{\tilde{f}} \setminus M_{\tilde{f}}(C)$.

Remark: The coins used to determine the neighbors in ground set $M_{\tilde{f}}$ of $C, \mathcal{H}_{\tilde{f}}(C)$ and $C', \mathcal{H}_{\tilde{f}}(C')$, $C \neq C'$ are completely independent

Note that the answer-vertices in the same class will have the same neighbors in $M_{\tilde{f}}$. A vertex $v_2 \in M_{\tilde{f}}$ is connected to either all the answer-vertices in a class or to none of them. See an example in Figure 1.4.

The final random graph $\hat{G}(I)$: The final $\hat{G} = (\hat{V}(I), \hat{E}(I))$ graph is derived from $\tilde{G}$ by the following modifications:

1. To get $\hat{V}_1$ duplicate every vertex in $v \in \tilde{V}_1$ $|\tilde{V}_2|/6^\ell$ times (to duplicate a vertex v i times is to replace v by i copies of v all with the same neighbors as v). Here we use the fact that 6^ℓ divides m, and so $|\tilde{V}_2|/6^\ell$ is an integer. After the modification, the size of every $A_{\overline{e}} B_{\overline{e'}}$ exactly equals $|\tilde{V}_2|$.

2. $\hat{V}_2 \leftarrow \tilde{V}_2$, namely, $\tilde{V}_2$ is not changed

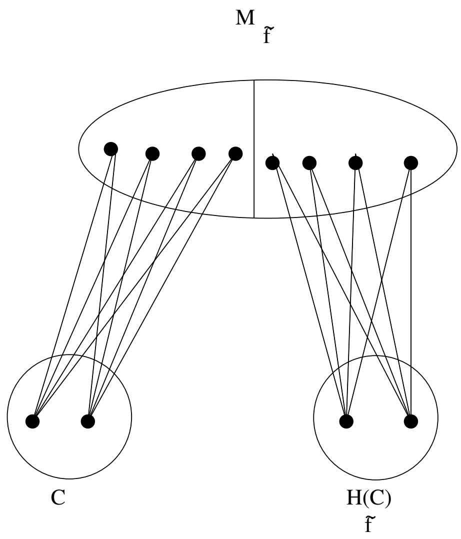

Fig. 2: The sets C and $H_{\tilde{f}}(C)$ have complementing neighbors in $M_{\tilde{f}}$. In this example $\ell = 1$ and $m = 8$.

3. Clique $A_{\overline{e}}, B_{\overline{e'}}$: For every $\overline{e}$ and $A_{\overline{e}}$ add an edge between all pairs $u, v \in A_{\overline{e}}$. Similarly, for for every $\overline{e'}$ and $B_{\overline{e'}}$ and add an edge between all pairs $z, w \in B_{\overline{e'}}$.
4. In addition, clique the vertices of $\hat{V}_2$, namely, add an edge between every two vertices of $\hat{V}_2$.

Since the number of variables is large, the following table gives a summary of symbols, their sizes, and their meaning.

2 Intuition and technical claims

2.1 Trapdoor: a large grundy number of a yes-instance

The construction is best explained via a yes-instance. Intuitively, in a yes-instance we can select $|\hat{V}_2| = |\hat{V}_2|$ subsets of $\hat{V}_1$ that are dominating sets of $\hat{V}_2$. These dominating sets form the first half of the color-classes in the order. The next half of the color-classes after the dominating sets are the vertices of $\hat{V}_2$, each forming its own color-class.

Formally:

Lemma 2.1 *The graph $\hat{G}$ that corresponds to a yes-instance I admits a grundy number of size* $2 \cdot |\hat{V}_2|$

Proof: Let I be a yes-instance. Choose a collection $\{\mathcal{S}_i\}_{i=1}^{|\hat{V}_2|}$ of subsets of $\hat{V}_1$ so that each $\mathcal{S}_i$ contains exactly one vertex from $A_{\overline{e}}$, $B_{\overline{e'}}$ and corresponds to a min-rep cover. The existence of such $\mathcal{S}_i$ follows from Lemma 1.6 and because every $\tilde{V}_1$ vertices is duplicated $|\hat{V}_2|/6^{\ell}$ times. We place these sets as the first $|\hat{V}_2|$ color-classes in the ordering. After that, we place $|\hat{V}_2|$ color-classes, each containing a single $|\hat{V}_2|$ vertex.

Symbol	**Description**	**Size**
$I = (\mathcal{V}(I), \mathcal{E}(I))$	The graph with the $3-$coloring gap	$\lvert V(I)\rvert = n.$
ℓ	The number of parallel repetitions	ℓ is a large constant.
$\mathsf{MRG}(I) = (U_1, U_2, E)$	The min-rep graph	$\lvert U_i\rvert = q \cdot 6^\ell = (15 \cdot n)^\ell.$
q	The number of questions for each prover	$q = n^\ell \cdot 5^\ell / 2^\ell.$
$SG(I)$	The supergraph with vertex set $\{A_{\overline{e}}\}, \{B_{\overline{e'}}\}$	$\lvert\{A_{\overline{e}}\}\rvert = \lvert\{B_{\overline{e'}}\}\rvert = q.$
d	The graph $SG(I)$ is d-regular	$d = 10^\ell.$
h	The number of superedges	$h = d \cdot q = (25 \cdot n)^\ell.$
$\tilde{G} = (\tilde{V}_1, \tilde{V}_2)$	Intermediate random graph. $\tilde{V}_1 = U_1 \cup U_2$	$\lvert\tilde{V}_1\rvert = 2 \cdot (15 \cdot n)^\ell.$
$M_{\tilde{f}}$	A ground set	$\lvert M_{\tilde{f}}\rvert = m \geq n^{2 \cdot \ell},\ 6^\ell \backslash m$
$\tilde{V}_2 = \hat{V}_2$	$\tilde{V}_2 = \hat{V}_2 = \bigcup_{\tilde{f} \text{ is a superedge}} M_{\tilde{f}}.$	$\lvert\tilde{V}_2\rvert = \lvert\hat{V}_2\rvert = h \cdot m$
$\hat{G} = (\hat{V}_1, \hat{V}_2)$	Duplications of $\tilde{V}_1$ and cliqueing $\tilde{V}_2$, etc.	$\lvert A_{\overline{e}}\rvert = \lvert B_{\overline{e'}}\rvert = \lvert\hat{V}_2\rvert.$

Tab. 1: Summary of notation

We now show that the above is a legal partition. We first show that every $\mathcal{S}_i$ is a dominating set of $\hat{V}_2$. Observe that every $\mathcal{S}_i$ contain a min-rep cover of every superedge $\tilde{f} = (A_{\overline{e}}, B_{\overline{e'}})$. Namely, there are $a \in \mathcal{S}_i \cap A_{\overline{e}}$ and $b \in \mathcal{S}_i \cap B_{\overline{e'}}$ so that $(a, b) \in E$.

It follows that a, b are adjacent to complementing halves of the ground set $M_{\tilde{f}}$, and thus are a dominating set of $M_{\tilde{f}}$. As this is true for every superedge, $\mathcal{S}_i$ is a dominating set of $\hat{V}_2$.

In addition, each $\mathcal{S}_i$ is an independent set; Recall that for every $\overline{e}$,$\overline{e'}$ $A_{\overline{e}}, B_{\overline{e'}}$ are cliques, but other than that there are no edges inside $\hat{V}_1$. As no $\mathcal{S}_i$ color-class contains two or more vertices of the same $A_{\overline{e}}$ or the same $B_{\overline{e'}}$, the claim follows.

Finally, we show that every vertex v in some $\mathcal{S}_i$ has a neighbor in every $\mathcal{S}_j$, for $j < i$. Say without loss of generality that $v \in A_{\overline{e}}$. Observe that $\mathcal{S}_j$ contains exactly one other vertex, say u, from $A_{\overline{e}}$. In addition, by construction $A_{\overline{e}}$ is a clique. The required claim follows, and we have shown that the above is a legal size $2 \cdot |\hat{V}_2|$ partition. $\square$

2.2 A technical claim for a no-instance: lucky pairs

For the rest of the paper we assume that I is a no-instance. Let $\Gamma(G)$ be the optimal grundy number for I.

We later define a universal constant ρ (it will turn out later that $\rho = 3$ is enough for our purposes) and another universal constant $\mu = \mu(\rho)$ that depends only on ρ.

In order to show that our reduction preserves the gap between a yes and a no-instance of $I = (\mathcal{V}(I), \mathcal{E}(I))$ we need to show that for a no-instance the grundy number cannot be arbitrarily close to $2 \cdot |\hat{V}_2|$. More precisely, that there exists some universal constant μ so that *opt* $\leq (2 - \mu)||\hat{V}_2|$. For the sake of contradiction for the rest of the paper we assume that for every universal constant μ, $\Gamma(G) = \{C_1, C_2, \ldots, C_{opt}\}$ with *opt* $\geq (2 - \mu)||\hat{V}_2|$ (*opt* is the number of color-classes in $\Gamma(G)$).

Assumption 2.1 *I is a no-instance and opt* $\geq (2 - \mu)||\hat{V}_2|$

For the rest of the section, let $\tilde{f} = (A_{\overline{e}}, B_{\overline{e'}})$ be a superedge.

Intuitively, our aim is to show that the color-classes for a no-instance must be “large.” Hence, any grundy partition must have small size. The no-instance cannot “coordinate efforts” by using consistent answer-vertices, and so the connection between the answer-vertices and the $M_{\tilde{f}}$ ground sets are essentially random. Intuitively we show that for any superedge $\tilde{f} = (A_{\overline{e}}, B_{\overline{e'}})$ a subset $C' \subseteq \hat{V}_1$ with only few vertices of $A_{\overline{e}} \cup B_{\overline{e'}}$ cannot dominate a "large" subset $M' \subseteq M_{\tilde{f}}$ unless C' min-rep covers $\tilde{f}$. We now define these terms precisely. Let $M' \subseteq M_{\tilde{f}}$ and $C' \subseteq \hat{V}_1$. Let ρ be some universal constant. Let μ be a universal constant that depends only on ρ. Let $DS(C', M')$ be the events that C' is a dominating set of M'; namely, each vertex in M' has a neighbor in C'.

Definition 2.1 Lucky Pairs
*The pair (C', M') is a $(\rho, \mu)-$*lucky pair *(or just lucky pair for short) if the following properties hold:*

1. $|M'| \geq (1 - 4 \cdot \mu)m$
2. *C' does not min-rep covers $\tilde{f}$ (namely, there are no $a \in C' \cap A_{\overline{e}}$, $b \in C' \cap B_{\overline{e'}}$ so that $(a, b) \in E$).*
3. *$DS(C', M')$ holds*
4. $|(A_{\overline{e}} \cup B_{\overline{e'}})| \leq 8 \cdot \rho$.

We later see that with high probability for any I (I can be a yes or a no-instance) there are no lucky pairs in the graph . Further, we shall show that if no lucky pairs exist, a no-instance must have a small size grundy number.

First, compute the probability that a given fixed triplet $\tilde{f}, M', C'$ as in Definition 2.1 is a lucky pair.

Claim 2.2

$$Pr(DS(C', M')) \leq exp(-\Omega(m/exp(\Theta(\rho)))).$$

Proof: Since C' does not min-rep covers $\tilde{f}$, it follows that every answer-vertex in $C' \cap (A_{\overline{e}} \cup B_{\overline{e'}})$ is connected to a random half of ground set $M_{\tilde{f}}$ (recall that only $A_{\overline{e}} \cup B_{\overline{e'}}$ answer-vertices are connected to $M_{\tilde{f}}$). The neighbors in $M_{\tilde{f}}$ of every pair of $x_1, x_2 \in C'$ are completely independent over the probability space.

The probability that $v_2 \in M'$ is not dominated by C' (does not have a neighbor in C') is $1/2^{8\rho}$. The expected number of elements in M' not dominated by C' (elements having no neighbor in C') is $\nu = |M'|/2^{8\rho}$. Let X be the number of not dominated by C' elements in M'. Let $\delta = 1 - 1/\nu$. Using the Chernoff bound (cf., **?**) we get that:

$$Pr(X = 0) = Pr(X < 1) = Pr(X < (1 - \delta)\nu) < exp(-\nu\delta^2/2) = 1/exp(\Omega(m/exp(\Theta(\rho)))).$$

□

We now upper bound the probability that *there exists some* lucky pair. Namely, we have computed the required probability for a fix $3-$tuple $\tilde{f}$ and M' and C'. We now bound the probability that such an event occurs over all possible $3-$tuples.

Claim 2.3 *For every universal constant ρ, we can choose μ to be a small enough universal constant depending on ρ so that the probability that there exists a lucky pair is $exp(-\Omega(m))$.*

Proof: Recall that we started with the intermediate random graph $\tilde{G}(I)$ and then duplicate every vertex in $\tilde{V}_1$ $|\hat{V}_2|/6^\ell$ times to get $\hat{V}_1$. If we remove duplicates from C', the event $DS(C', M')$ is not affected (the removed copies dominate the same vertex set as the copy that is not discarded). Thus, we may assume that $C' \subseteq \tilde{V}_1$, and the probability of $DS(C', M')$ can be computed over $\tilde{G}$, the graph before the duplications.

We use the union bound, namely, bound the probability over all $3-$tuples by the sum of probabilities for every individual tuple. By symmetry, this equals the number of $3-$tuples, times the probability derived in Claim 2.2.

Hence, we next bound from above the number of $\tilde{f}, C', M'$ triplets over $\tilde{G}$.

1. The number of $C' \subset \tilde{V}_1$ sets: Recall that $q = |\mathcal{E}(I)^\ell| = (\frac{5}{2})^\ell |\mathcal{V}(I)|^\ell$. The number of vertices in $\tilde{G}$ is
$$2 \cdot q \cdot 6^\ell = 2 \cdot (15 \cdot n)^\ell = O(15^\ell \cdot n^\ell).$$
The number of subsets C' of $\tilde{V}_1$ is thus $exp(\Theta(15^\ell \cdot n^\ell))$.

2. There are at most $h = d \cdot q = (25 \cdot n)^\ell$ superedges

3. The number of M' sets: The number of subsets of size $(1 - 4 \cdot \mu)m$ of $M_{\tilde{f}}$ is bounded as follows: We use,
$$\binom{m}{\alpha m} \leq 2^{mH(\alpha)}$$
with $H(\alpha)$ is the binary entropy function that goes to 0 when α goes to 0 (see **?**). Thus, the number of $(1 - 4 \cdot \mu)m$ size subsets of of $M_{\tilde{f}}$ is bounded by
$$\binom{m}{4 \cdot \mu m} \leq 2^{mH(4\mu)}.$$

Multiplying the four upper bounds, we get a bound of at most
$$exp(\Theta(15^\ell \cdot n^\ell)) \cdot 25^\ell \cdot n^\ell \cdot 2^{mH(4\mu)}$$
$3-$tuples.

Thus, the probability that there exist any lucky pair is at most
$$\frac{exp(\Theta((15n)^\ell)) \cdot (25 \cdot n)^\ell \cdot 2^{mH(4\mu)}}{exp(\Omega(m/exp(\Theta(\rho))))}.$$

Since $m \geq n^{2\ell}$, taking μ small enough, the above probability is exponentially small in m. The claim follows. $\square$

Claim 2.3 shows that a subset C' has to be "large" in order to cover "many" vertices of $M_{\tilde{f}}$. This is used in the next section to show that a grundy number for a no-instance cannot be too large.

3 Analysis of the reduction

We based the analysis on the following lemma:

Lemma 3.1 *Under Assumption 2.1, a lucky pair must exist.*

First, note that once Lemma 3.1 is proved, we are done:

Corollary 3.2 *If I is a no-instance, then for small enough μ, with probability* $1 - exp(-\Theta(m))$ *$opt \leq (2-\mu)|\hat{V}_2|$.*

Proof: If I is a no-instance, then from Claim 2.3 probability at least $1-exp(-\Theta(m))$ there is not a lucky pair. Hence by Lemma 3.1 Assumption 2.1 cannot hold, and so $opt \leq (2-\mu)|\hat{V}_2|$ □

Theorem 3.3 *There exists a constant τ so that approximating the grundy number problem within $(1+\tau)$ is NP-hard, under randomized reductions*

Proof: By Lemma 2.1, $\hat{G}$ that corresponds to a yes-instance has grundy number bounded below by $2|\hat{V}_2|$. By Corollary 3.2, with high probability a no-instance has a size $(2-\mu)|\hat{V}_2|$ size partition for some universal constant μ. The claim follows. □

The rest of the paper is devoted to proving Lemma 3.1.

The first thing we need to do is identify color-classes C_i in $\Gamma(G)$ that are fully contained in $\hat{V}_1$, namely, do not contain $\hat{V}_2$ vertices (they will play the role of C' from the definition of a lucky pair) followed by many vertices of the same $M_{\tilde{f}}$ set. These vertices will become M'. Since C_i is before M' in the ordering, C_i has to be a dominating set of M', and the situation mimics Claim 2.3.

We split its color-classes in $\Gamma(G)$ into three types:

1. **Type 1:** Color-classes containing only $\hat{V}_1$ vertices.
2. **Type 2:** Color-classes containing only $\hat{V}_2$ vertices (because $\hat{V}_2$ is a clique, such a cluster contains a single $\hat{V}_2$ vertex).
3. **Type 3:** Color-classes containing at least one vertex of $\hat{V}_1$ and exactly one vertex of $\hat{V}_2$.

Let $I_i, 1 \leq i \leq 3$, be the number of color-classes of types $1, 2, 3$, respectively. We now look for color-classes not containing $\hat{V}_1$ vertices. Let i_{max} be the highest index so that $C_{i_{max}}$ contains a $\hat{V}_1$ vertex (namely, $C_{i_{max}}$ is a type 1 or 3 cluster). Then:

Claim 3.4 *If $n = |\mathcal{V}(I)|$ is large enough, then*

$$i_{max} \leq (1+\mu)|\hat{V}_2|.$$

Proof: Let $v \in \hat{V}_1 \cap C_{i_{max}}$. Note that each $C_j,\ j < i_{max}$ must contain a neighbor of v. Recall that d is the degree of every supervertex in the supergraph $SG(I)$. The degree of v in $\hat{G}$ is less than $|\hat{V}_2| + d \cdot m$. Indeed, say without loss of generality that $v \in A_{\overline{e}}$. Thus, v has $|\hat{V}_2| - 1$ neighbors in $A_{\overline{e}}$ (recall that $A_{\overline{e}}$ was cliqued). In addition, for every one of the d superedges involving $A_{\overline{e}}$ in $SG(I)$, v is joined to half of the m vertices in the corresponding ground set.

It thus follows that $i_{max} < |\hat{V}_2| + d \cdot m = |\hat{V}_2| + |\hat{V}_2|/q$. Note that $q > n$. The claim follows. □

Corollary 3.5 *The last $(1-2\mu)|\hat{V}_2|$ color-classes in the optimal partition are type 2 color-classes, namely, are color-classes containing a single $\hat{V}_2$ vertex. In addition, the first $(1+\mu)|\hat{V}_2|$ color-classes include at least $(1-\mu)|\hat{V}_2|$ type 1 color-classes.*

Proof: The first part of the corollary follows directly from Claim 3.4. In addition, as the last $(1-2\mu)|\hat{V}_2|$ color-classes are of type 2 (by Claim 3.4), there could be at most $2\cdot\mu|\hat{V}_2|$ additional type 2 or type 3 color-classes among the first $(1+\mu)|\hat{V}_2|$ color-classes. Indeed, each type 2 or 3 cluster contains a vertex of $\hat{V}_2$. The claim follows. □

Corollary 3.6 *The optimum solution has $(1-\mu)|\hat{V}_2|$ color-classes that are type 1 (subsets of $\hat{V}_1$) and are followed by $(1-2\mu)|\hat{V}_2|$ type 2 color-classes (color-classes containing a single $\hat{V}_2$ vertex).*

We focus on this subset of color-classes:

Definition 3.1 *The $(1-\mu)|\hat{V}_2|$ (at least) color-classes of type 1 that are followed by $(1-2\cdot\mu)|\hat{V}_2|$ (at least) type 2 color-classes are the* important *color-classes.*

The important type 1 color-classes C_i will play the role of C' in the definition of a lucky pair. However, toward showing the existence of a lucky pair, we need to show that many $M_{\tilde{f}}$ ground sets contribute at least $(1-4\mu)$ vertices to important type 2 color-classes. Let $\mathcal{U}\subseteq\hat{V}_2$ be the set of all $\hat{V}_2$ vertices appearing in type 2 important color-classes. Exactly one $\hat{V}_2$ vertex appears in each color-class.

Definition 3.2 *Let $M'=M_{\tilde{f}}\cap\mathcal{U}$. We say that $M_{\tilde{f}}$ is* proper *if $|M'|=|M_{\tilde{f}}\cap\mathcal{U}|\geq(1-4\mu)\cdot m$*

Claim 3.7 *At least half of the ground sets (namely, $h/2$ of the ground sets) contribute at least $(1-4\cdot\mu)m$ elements to $\mathcal{U}$.*

Proof: Say that i of the ground sets contribute less than $(1-4\cdot\mu)m$ to $\mathcal{U}$, and the rest of the $h-i$ ground sets contribute to $\mathcal{U}$ more than that. We now bound from above the size of $\mathcal{U}$. Observe that by Corollary 3.6, $|\mathcal{U}|\geq|\hat{V}_2|-2\cdot\mu|\hat{V}_2|$. On the other hand, by the definition of i:

$$|\mathcal{U}|\leq i(1-4\cdot\mu)m+(h-i)m.$$

Since $(1-2\mu)|\hat{V}_2|\leq|\mathcal{U}|$ and $h\cdot m=|\hat{V}_2|$, we get that:

$$-2\cdot\mu|\hat{V}_2|\leq-4\cdot\mu\cdot i\cdot m.$$

So, $i\leq h/2$ and $h-i\geq h/2$. Hence, at least $h-i\geq h/2$ ground sets contribute more than $(1-4\cdot\mu)m$ elements to $\mathcal{U}$. □

In Claim 2.3, C' does not min-rep covers M'. We now show that this relation also holds for many important color-classes pairs: C_i an important color-class of type 1, and M', with $M'=M_{\tilde{f}}\cap\mathcal{U}$, $M_{\tilde{f}}$ a proper ground set.

Consider important color-class C_i of type 1. Call C_i *problematic* if it min-rep covers at least $h/4$ of the proper superedges (namely, for roughly half of the proper superedges $\tilde{f}=(A_{\overline{e}},B_{\overline{e'}})$, C_i contains at least one pair $(a,b)\in E$, $a\in A_{\overline{e}}$, $b_j\in B_{\overline{e'}}$ and $(a,b)\in MRG(I)$). Since I is a no-instance, using Lemma 1.6 we show that the number of problematic C_i is small.

Claim 3.8 *Given a constant μ, we can choose the parallel repetition ℓ to be a large enough universal constant (depending on μ) so that every problematic color-class C_i contains at least $2q/\mu$ vertices.*

Proof: As C_i min-rep covers $\alpha = 1/4 > 1/5$ fraction of the edges, it follows from Lemma 1.6 that, if we put $|C_i| \geq 2qk$, then k must satisfy:

$$1/5 < \left(32 \cdot k^2 \cdot 2^{-c\ell}\right)^{1/3}.$$

The claim follows for a large enough ℓ. □

As $\hat{V}_1 = 2q \cdot |\hat{V}_2|$, we get the following corollary.

Corollary 3.9 *The number of type 1 problematic C_i is bounded by $\mu \cdot |\hat{V}_2|$. Hence, there are at least $(1-2\mu)|\hat{V}_2|$ type 1 non-problematic important color-classes C_i.*

Proof: The total number of $\hat{V}_1$ vertices is $2q \cdot |\hat{V}_2|$. If C_i is not problematic, C_i contains at least $2q/\mu$ elements (Claim 3.8). Let j be the number of type 1 problematic color-classes. Clearly:

$$j \cdot 2q/\mu \leq |\hat{V}_1| = 2q \cdot |\hat{V}_2|.$$

Thus,

$$j \leq \mu \cdot |\hat{V}_2|.$$

As from Corollary 3.6 we know that there are at least $(1-\mu)|\hat{V}_2|$ important color-classes, the claim follows. □

Finally, we prove Lemma 3.1. Namely, that under Assumption 2.1 a lucky pair exist.

Proof: We show that if a lucky pair does not exist, then for μ small enough, $|C_i| \geq 3q$. This is a contradiction: C_i has at most one vertex per $A_{\overline{e}}$ and $B_{\overline{e'}}$, $|C_i| \leq 2q$.

Let C_i be an arbitrary important and non-problematic type 1 color-class. We know from Claim 3.7 that at least $h/2$ $M_{\tilde{f}}$ contribute to $\mathcal{U}$ at least $(1-4\cdot\mu)m$ vertices. For every $\tilde{f}$ $M' = M_{\tilde{f}} \cap \mathcal{U}$ has size $|M'| \geq (1-4\cdot\mu)m$, $M' \subseteq \mathcal{U}$. Among those, at least $h/2-h/4$ are $M_{\tilde{f}}$ so that C_i does not min-rep covers $\tilde{f}$. All these $h/4$ M' sets contain vertices that appear after C_i in the partition. Thus, C_i is a dominating set of all those M'.

Choose $\rho = 3$. Choose μ to be the appropriate universal constant (that follows by the choice of $\rho = 3$). Examine the pair $C' = C_i$ and $M' = M'(\tilde{f})$ as defined above. The properties we already know are:

1. $|M'| \geq (1-4\mu)m$.
2. C_i does not min-rep cover $\tilde{f}$.
3. $DS(C', M')$ holds (C' is a dominating set of M').

The only missing property from the definition of a lucky pair is $|C' \cap (A_{\overline{e}} \cup B_{\overline{e'}})| \leq 8 \cdot \rho$.

By the assumption that there is no lucky pair, $|C_i \cap (A_{\overline{e}} \cup B_{\overline{e'}})| > 8 \cdot \rho = 24$

We now bound from below the size of C_i. Every one of the $h/4$ (or more) proper superedges $\tilde{f}$ not min-rep covered by C_i "adds" to C_i at least 24 elements of its set $A_{\overline{e}} \cup B_{\overline{e'}}$ for a total of $24 \cdot h$. However, in this way, every $A_{\overline{e}}$ and every $B_{\overline{e'}}$ are counted d times. Thus:

$$|C_i| \geq \frac{24 \cdot h}{4 \cdot 2d} = 3 \cdot q.$$

This gives the desired contradiction, and hence a lucky pair must exist. □

4 Open problems

It would be interesting to get a super-constant lower bound on the approximability of the grundy number problem. In addition, no nontrivial approximation algorithm for the grundy number problem exists even on bipartite graphs. In the author's opinion, any $o(n)$ approximation is of interest. See **?** for an $O(n^{4/5})$ approximation algorithm for the achromatic number of bipartite graphs. The hope would be that **?** can be used to a sublinear approximation for grundy number of bipartite graphs. However, there seem to be considerable difficulties involved.

Acknowledgment

The author thanks Rajiv Gandhi and Magnus M. Halldórsson for useful discussions and comments.

References