Graph Theoretic Approach to Parallel Gene Assembly

Tero Harju
Department of Mathematics, University of Turku,
Turku Centre for Computer Science,
FIN-20014 Turku, Finland
harju@utu.fi

Chang Li
Department of IT, Abo Akademi University,
Turku Centre for Computer Science,
FIN-20520 Turku, Finland
lchang@abo.fi

Ion Petre
Academy of Finland and
Department of IT, Abo Akademi University,
Turku Centre for Computer Science,
FIN-20520 Turku, Finland
ipetre@abo.fi

August 23, 2007

Abstract
We study parallel complexity of signed graphs motivated by the highly complex genetic recombination processes in ciliates. The molecular gene assembly operations have been modeled by operations of signed graphs, i.e., graphs where the edges have a sign + or −. In the optimization problem for signed graphs one wishes to find the parallel complexity by which the graph can be reduced to the empty graph. We relate parallel complexity to matchings in graphs for some natural graph classes, especially bipartite graphs. It is shown, for instance, that a bipartite graph \( G \) has parallel complexity one if and only if \( G \) has a unique perfect matching. We also formulate some open problems of this research topic.

Keywords. Gene assembly, signed graphs, parallel assembly, local complement, split graphs, double-split graphs, perfect matching

1 Introduction

In this paper we take a graph theoretic approach to gene assembly of ciliates. After shortly introducing the biological motivation of the topic, we study the
model of gene assembly in terms of signed graphs and their operations. Our main
problems concern low parallel complexity of these operations on signed graphs
which are undirected graphs with edges having a sign + or −. We consider
three basic operations (negative, positive and double rules) on signed graphs
that correspond to the gene assembly operations in the genetic transformation
process in ciliates. Each signed graph can be reduced to the empty graph
using a sequence of these operations. However, some of the operations may be
applicable simultaneously to a signed graph, which raises the question of parallel
complexity of the graphs: how many parallel steps are needed in the reduction
to the empty graph.

We relate parallel complexity to matchings in graphs for some natural graph
classes, especially bipartite graphs. It is shown, for instance, that a (unsigned)
bipartite graph $G$ can be reduced to the empty graph in one parallel step if
and only if $G$ has a unique perfect matching. We also show that all trees with
positive signs on the edges, can be reduced in one step to a graph with positive
signs on the edges.

A visible part of the present paper consists of open problems of this research
topic.

2 Preliminaries

2.1 Biological motivation

Ciliates are ancient, eucaryotic single-cell organisms [2], that inhabit practically
all types of aqueous environments. The most striking feature, common to all
species of ciliates, is that they possess two dissimilar nuclei, micronucleus and
macronucleus, which have different functions in the cell. The macronucleus
comprises the household genome carrying the RNA transcription needed for cell
growth and proliferation [17]. It consists of short gene-size chromosomes [18],
and it is generated from the micronucleus after cell mating. The micronucleus,
on the other hand, is active during cell mating when DNA material is exchanged
between two cells. It has no known function in the normal life of the cell.
The micronuclear genes are located in long chromosomes where each gene can
be fragmented in several segments known as macronuclear destined sequences
(MDSs). The MDSs are separated by, often quite long, non-coding segments
(internally eliminated segments, IESs), and, moreover, the MDSs may occur in
permuted order and some of them even inverted.

The process of converting micronucleus to macronucleus involves massive
DNA rearrangement with respect to the genomic difference of the two nu-
clei [13]. During the macronucleus development all spacer sequences (IESs)
between MDSs are excised and the remaining MDSs are recombined in correct
order to form the genes. This process is referred as gene assembly.

From computational point of view, MDSs are arranged as the linked list data
structure, where the outgoing pointer of one data matches with the incoming
pointer of the next data. Thus, the double occurred sequences at IES/MDS
junctions can be referred as pointers, and the whole process may be seen as
pointer-guided [7].

Concerning how IESs are recognized/excised and how MDSs are unscram-
bled/recombined, several models have been postulated [14, 15, 20]. In this paper,
we consider the intramolecular model [4], where three molecular operations are
postulated [19]: loop recombination, hairpin recombination and double-loop re-
combination. In each of these operations, the molecule folds on itself so that two
or more pointers get aligned and through recombination two or more MDSs get
combined into a bigger composite MDS. The process continues until all MDSs
have been assembled.

In the intramolecular model, the MDSs arrangement on a micronuclear gene
is formalized on three levels of abstraction [3]: MDS descriptors (denoting the
sequence of the incoming pointers and outgoing pointers of MDSs), legal strings
(denoting the sequence and polarity of the double occurred pointers), and signed
graphs (denoting the overlap relation between the pointers and the pointer po-

larity). Correspondingly, the three molecular operations are formalized as the
reduction of MDSs pointers, strings and graphs.

In this paper we concentrate on the graph theoretic approach on gene as-

semble. The basic problem is to find parallel gene assembly strategies for a
given signed graph associated with a micronuclear gene pattern, and to give an
optimal parallel assembly strategy that requires the least assembly steps.

Our first problem is: when can multiple operations form a single step of
an assembly strategy? On molecular level, we can see that the MDSs whose
pointers do not occur on the overlapping repeats can be assembled in parallel.
On the graph level, we do not have a full characterization of applying multiple
assembly operations in parallel. The second problem is: in how many steps can
any given gene pattern be assembled? For signed graphs, this problem concerns
the upper bounds on the number of reduction steps. It turns out that this
bound does not completely corresponding to the number of pointers in the gene
patterns, but it depends on the type of graph and the reduction strategy been
applied.

In this paper, we investigate these two problems on various signed graphs,
e.g., bipartite and tripartite graph, trees, and also arbitrary graphs. For some
types of graphs we prove that the number of parallel reduction steps is bounded.

2.2 Signed graphs

In this section we introduce basic notions concerning signed graphs and the
reduction operations for signed graphs.

Let \( G = (V, E) \) be a finite undirected graph without self loops or multiple
edges. The number \(|V|\) of the vertices, is called the order of \( G \). For simplicity,
we write \( v \in G \) if \( v \) is a vertex, and \( e \in E(G) \), if \( e \) is an edge of \( G \). The
neighborhood of a vertex \( v \) is \( N_G(v) = \{ u \mid \{ v, u \} \in E(G) \} \). We call the set
\( N^+_G(v) = N_G(v) \cup \{ v \} \) the closed neighborhood of \( v \). We say that two vertices \( u, v \),
are twins if \( N^+_G(u) = N^+_G(v) \). The vertex \( v \) is isolated if \( N_G(v) = \emptyset \). We call \( G \)
discrete if all its vertices are isolated. A subset \( A \subseteq G \) is stable, if there are no
edges \( \{ v, u \} \) with \( v, u \in A \). A graph \( G \) is called a clique (or a complete graph)
if any two vertices in \( G \) are adjacent.

A subset \( M \subseteq E(G) \) is a matching of a graph \( G \) if no two edges in \( M \) have
a common end point. A matching \( M \) is said to be perfect, if each vertex is
an endpoint of an edge in \( M \). A matching \( M \) is maximal if \( M \cup \{ e \} \) is not a
matching for all edges \( e \notin M \), and \( M \) is maximum matching, if its cardinality
is the largest possible.
A signed graph $G = (V, E, \sigma)$ consists of a graph $(V, E)$ together with a labeling function $\sigma: V \to \{+, -\}$ of the vertices. (A signed graph is often called a marked graph in literature.) A vertex $v \in G$ is said to be positive (negative, resp.) if $\sigma(v) = +$ ($\sigma(v) = -$, resp.) We let

$$G^+ = \{v \mid \sigma(v) = +\} \quad \text{and} \quad G^- = \{v \mid \sigma(v) = -\}$$

the corresponding sets of positive and negative vertices in $G$. We say that a signed graph is a negative graph (a positive graph, resp.) if all its vertices are negative (positive, resp.) Also, an edge $e = \{v, u\}$ is called negative, if $v, u \in G^-$. Let $G = (V, E, \sigma)$ be a signed graph. For each $A \subseteq G$, let $G - A$ denote the subgraph induced by the set of vertices $V \setminus A$.

Denote by $\text{loc}_c(G)$ the local complement of $G$ at $v$: $\text{loc}_c(G) = (V, E', \sigma')$, where

\[
{x, y} \in E' \iff \begin{cases} 
{x, y} \notin E \text{ and } x, y \in N_G(v), & \text{or} \\
{x, y} \in E \text{ and, } x \notin N_G(v) \text{ or } y \notin N_G(v).
\end{cases}
\]

Also, $\sigma'(x) = +$ if and only if $\sigma(x) = -$, for all $x \in N_G(v)$ and $\sigma'(x) = \sigma(x)$, otherwise.

A split graph is a graph $G$ in which the vertex set can be partitioned into two sets $K$ and $A$ such that $K$ induces a clique and $A$ induces a stable set in $G$.

We say that $G = (V, E)$ is a double split graph if the vertex set can be partitioned into four sets $A = \{a_1, \ldots, a_m\}, B = \{b_1, \ldots, b_m\}, C = \{c_1, \ldots, c_n\}, D = \{d_1, \ldots, d_n\}$ for some $m, n \geq 2$ such that:

- $A$ and $B$ are stable sets in $G$;  
- $C$ and $D$ are cliques in $G$;
- $\{a_i, b_i\} \in E$ for $1 \leq i \leq m$ and $\{c_j, d_j\} \notin E$ for $1 \leq j \leq n$;
- $\{a_i, b_j\} \notin E$ for $i \neq j$ and $\{c_i, d_j\} \in E$ for $i \neq j$;
- for each $1 \leq i \leq m$ and $1 \leq j \leq n$, either $\{a_i, c_j\} \in E$ and $\{b_i, d_j\} \in E$ or $\{a_i, d_j\} \in E$ and $\{b_i, c_j\} \in E$.

For two signed graphs $G$ and $H$ on disjoint vertex sets, say $V(G)$ and $V(H)$, let $G \oplus H$ be their disjoint union, i.e., the vertex set of $G \oplus H$ is $V(G) \cup V(H)$, and its edges form the set

$$E(G \oplus H) = E(G) \cup E(H).$$

The complete connection $G \otimes H$ has the vertex set $V(G) \cup V(H)$ and the edge set is

$$E(G \otimes H) = E(G) \cup E(H) \cup \{\{v, u\} \mid v \in G, u \in H\}.$$

### 2.3 Graph reduction rules

The three molecular operations are formalized as reduction rules for signed graphs as followed. For details how the molecular operations are translated to these formal operations, we refer to [3].

Let $v$ and $u$ be two vertices.
The \textit{(graph) negative rule} for $v$ is applicable to $G$ if $v \in G^-$ and it is isolated in $G$. The result is the signed graph $\text{gnr}_v(G) = G - v$. The domain of $\text{gnr}_v$ is $\text{dom}(\text{gnr}_v) = \{v\}$.

The \textit{(graph) positive rule} for $v$ is applicable to $G$ if $v \in G^+$. The result is the signed graph $\text{gpr}_v(G) = \text{loc}_v(G) - \{v\}$. The domain of $\text{gpr}_v$ is $\text{dom}(\text{gpr}_v) = \{v\}$.

The \textit{(graph) double rule} for $v \neq u$ is applicable to $G$ if $v, u \in G^-$ and $e = \{v, u\} \in E(G)$. The result is the signed graph $\text{gdr}_e(G) = (G\setminus\{v, u\}, E', \sigma')$ where $\sigma'$ equals $\sigma$ restricted to $G - \{v, u\}$, and $E'$ is obtained from $E$ by complementing the edges that join vertices of $N_G(v)$ and $N_G(u)$. This means that the status of a pair $\{x, y\}$ as an edge will change if and only if $v$ and $u$ belong to different sets $N_G(v) \setminus N_G(u)$, $N_G(u) \setminus N_G(v)$, $N_G(v) \cap N_G(u)$.

The domain of $\text{gdr}_e$ is $\text{dom}(\text{gdr}_e) = \{v, u\}$.

We let $\text{Op} = \{\text{gnr}_*, \text{gpr}_*, \text{gdr}_{e,u} | v \neq u\}$ denote the full set of the operations. Let $G$ be a signed graph. A composition $\varphi = \varphi_n \circ \cdots \circ \varphi_1$ of operations $\varphi_1, \ldots, \varphi_n \in \text{Op}$ is called a graph reduction for $G$ if $\varphi$ is applicable to $G$. We say that $\varphi$ is a successful reduction strategy for $G$ if $\varphi(G) = \emptyset$.

**Example 1.** Consider the signed graph $G$ of six vertices shown in Figure 1(a).

One successful reduction strategy of $G$ is: $\varphi = \text{gdr}_{4,5} \circ \text{gpr}_6 \circ \text{gnr}_7 \circ \text{gnr}_3 \circ \text{gnr}_2$, see Figure 1(b) and (c).

![Figure 1](image)

Figure 1: (a) A signed graph $G$. (b) Applying $\text{gnr}_2$, $\text{gnr}_3$ and $\text{gnr}_7$ to $G$, reduces all the isolated negative vertices. (c) Applying $\text{gpr}_6$ reduces $G$ to a negative edge of $\{4, 5\}$, where $\text{gdr}_{4,5}$ is applicable.

\section{Parallelism}

Intuitively, a set of operations can be applied in parallel if and only if each application is independent of the others. In other words, a number of operations can be applied in parallel to a signed graph if they can be (sequentially) applied in any order to that graph. Note that this is consistent with how parallelism and concurrency are defined in Computer Science. The following gives the definition of parallel application of the three molecular operations on a signed graph.

**Definition 1.** Let $S \subseteq \text{Op}$ be a set of $k$ rules and let $G$ be a signed graph. We say that the rules in $S$ can be applied in parallel to $G$ if for any ordering $\varphi_1, \varphi_2, \ldots, \varphi_k$ of $S$, the composition $\varphi_k \circ \cdots \circ \varphi_1$ is applicable to $G$. 

5
Based on the definition of parallelism, which presumes that the rules are applicable in any possible order, the following theorem shows that the result is always the same regardless of the order in which they are applied.

**Theorem 3.1** ([8]). Let $G$ be a signed graph and let $S \subseteq \text{Op}$ be a set of operations applicable in parallel to $G$. Then for any two compositions $\varphi$ and $\psi$ of the operations of $S$, $\varphi(G) = \psi(G)$.

Based on Theorem 3.1, we can write $S(G) = \varphi(G)$ for any set $S$ of operations applicable in parallel to $G$ and any composition $\varphi$ of these operations. We define the parallel complexity as follows.

**Definition 2** ([10]). Let $G$ be a signed graph, and let $S_1, \ldots, S_k \subseteq \text{Op}$ be sets of operations each applicable in parallel to $G$. If $(S_k \circ \ldots \circ S_1)(G) = \emptyset$, then we say that $S = S_k \circ \ldots \circ S_1$ is a parallel reduction strategy for $G$. In this case the parallel complexity of $S$ is $c(S) = k$. The parallel complexity of the signed graph $G$ is:

$$c(G) = \min \{c(S) \mid S \text{ is a parallel reduction strategy for } G\}.$$

We denote the **square** $C_4$ and **diamond** $D_4$ as the negative graphs shown in Figure 2. In both graphs, no any two edges can be reduced in parallel.

The following provides a simple criterium for two rules to be applicable in parallel.

**Theorem 3.2** ([9]). Let $G$ be a signed graph and let $\varphi, \psi \in \text{Op}$ be two rules applicable to $G$ such that $\text{dom}(\varphi) \cap \text{dom}(\psi) = \emptyset$.

(i) If $\varphi = \text{gnr}_v$ is a negative rule, then $\varphi$ and $\psi$ can be applied in parallel to $G$.

(ii) If $\varphi = \text{gpr}_v$ is a positive rule, then $\varphi$ and $\psi$ can be applied in parallel to $G$ if and only if $N_G(v) \cap \text{dom}(\psi) = \emptyset$.

(iii) Two double rules $\varphi$ and $\psi$ can applied in parallel to $G$ if and only if the subgraph of $G$ induced by $\text{dom}(\varphi) \cup \text{dom}(\psi)$ is not isomorphic to $C_4$ or $D_4$.

The problem of finding a maximum size set of double operations that can be applied in parallel to a given set of negative edges in a given signed graph $G$ remains open. This is related to the following problem.

**Problem 1.** For a given graph $G$ and a given set of negative edges $A$, decide efficiently whether $\text{gdr}_A = \{\text{gdr}_c \mid c \in A\}$ is applicable in parallel to $G$.
4 A bound on the parallel complexity

The main open problem of this study is whether the parallel complexity of signed graphs is finitely bounded, i.e., for each positive integer $n$, does there exist a signed graph with parallel complexity $\mathcal{C}(G) \geq n$?

For some types of graphs, an upper bound for parallel complexity is easy to obtain, the following are some of the examples.

Example 2. (a) For a discrete signed graph $G$, $\mathcal{C}(G) = 1$.

(b) For a signed clique $G$, $\mathcal{C}(G) \leq 2$. Indeed, if there exists a positive vertex $v \in G^+$, then $\text{gpr}_v(G)$ is discrete, and then one can apply in parallel the rules $\text{gpr}_u$ or $\text{gnr}_u$ for each remaining vertex $u$. On the other hand, if $G^+ = \emptyset$, then choose a maximal matching $M$ of the edges of the clique. Observe that, for any $e \in M$, the graph $\text{gdr}_e(G)$ is a smaller clique, and hence $M \setminus \{e\}$ is applicable to $\text{gdr}_e(G)$. It follows that $\text{gdr}_M = \{\text{gdr}_e \mid e \in A\}$ can be applied in parallel to $G$, and this results in the empty graph or a singleton graph depending whether the clique $G$ had even or odd order.

(c) Let $G$ be a negative complete bipartite graph, i.e., $G = S_1 \otimes S_2$ for two discrete graphs $S_1$ and $S_2$. Then $\mathcal{C}(G) \leq 2$. Indeed, for any $e \in E(G)$, $\text{gdr}_e(G)$ is either a discrete graph or it is empty (in case $G$ has exactly one edge).

The following lemma for double rules $\text{gdr}_e$ is needed in Theorem 4.4.

Lemma 4.1. Let $G$ be a signed graph. Then $G$ has no parallel applications of two double rules if and only if $G^+ = (H \otimes S_1) \oplus S_2$, where $H$ is a complete bipartite graph, and $S_1$ and $S_2$ are (possibly empty) discrete graphs.

Proof. We assume without loss of generality that the graph $G$ is negative.

For the direct implication, note that the result holds for graphs with up to 4 vertices. Especially, the square $C_4$ and the diamond $D_4$ can be represented in the required form, but, for instance the clique $K_4$ on four vertices cannot.

Consider a larger graph and note that in that graph there can be several isolated vertices, but it can have only one non-trivial connected component. We remove the isolated vertices and thus assume without loss of generality that the graph is connected.

We claim that for any $v, u \in G$, if $\{v, u\} \notin E(G)$ then $N(v) = N(u)$. For this, assume there exists a vertex $a \in N(v) \setminus N(u)$. Since $G$ is connected, there exists a vertex $b \in N(u)$. Now, the set $\{v, u, a, b\}$ induces a subgraph different from $C_4$ and $D_4$ and so, $\text{gdr}_{v,a}$ and $\text{gdr}_{u,b}$ are both applicable to $G$. This proves the present claim.

Let $S_1$ be a maximal stable subset of $G$. Hence each vertex $v \notin S_1$ is adjacent to a vertex in $S_1$, and, by the claim, $S_1$ and $W = G \setminus S_1$ are completely connected. Since $G$ has no cliques $K_4$ on four vertices, $W$ does not have triangles $K_4$ and thus no subgraphs $D_4$. Therefore all four element induced subgraphs are isomorphic to the square $C_4$. This is possible only if $W$ induces a complete bipartite subgraph $H$. Hence $G = H \otimes S_1$ as required.

The reverse implication is clearly true.
In a complete tripartite graph $H$ the vertex set is partitioned to three subsets $A_1, A_2, A_3$ and a pair $\{v, u\}$ is an edge if and only if $v \in A_i$ and $u \in A_j$ for $i \neq j$.

**Corollary 4.2.** A signed graph $G$ has no parallel applications of two double rules if and only if $G^{-}$ consists of a discrete graph and a complete tripartite graph, where some of the three components can also be empty. Moreover, in this case, if $G$ is negative, then $\mathcal{C}(G) \leq 2$.

**Proof.** In Lemma 4.1, $H \otimes S_1$ is a complete tripartite graph. For the last claim for negative $G$, applying $\text{gdr}_e$ for one edge $e \in E(G)$ results in a discrete graph, which reduces to the empty graph in one step. 

**Theorem 4.3.** Let $G$ be a negative graph with $n$ vertices. Then $\mathcal{C}(G) \leq n/4 + 2$.

**Proof.** If no double rule is applicable to $G$, then $G$ is discrete and thus $\mathcal{C}(G) = 1$. If at least two double rules $\text{gdr}_e$ and $\text{gdr}_f$ are applicable in parallel to $G$, then the order of $G$ is decreased by at least 4 in one step. Consider then the case, where exactly one $\text{gdr}$-rule is applicable to $G$. By Lemma 4.1, $G = H \otimes S$, with $H$ a complete tripartite graph and $S$ a discrete graph. Here $\mathcal{C}(G) = \mathcal{C}(H) \leq 2$, since for each edge $e \in H$, the graph $\text{gdr}_e(H)$ is discrete.

In the next we give a non-finite upper bound for signed graphs in general.

**Theorem 4.4.** Let $G$ be a signed graph with $n$ vertices. Then $\mathcal{C}(G) \leq n/2 + 4$.

**Proof.** We show that there is a reduction of each signed graph of $n \geq 4$ vertices which reduces the graph by two vertices in one step or four vertices in two steps.

Assume that in one step only one positive rule applies to $G$. In particular, $G$ must be connected. Since no double rule $\text{gdr}_e$ is applicable, the negative vertices $G^{-}$ form a stable set, and since at most one positive rule $\text{gpr}_e$ is applicable, the positive vertices $G^{+}$ form a clique.

Consider a positive vertex $v \in G^{+}$ of minimum degree $|N_G(v)|$, and let

$$A = G^{+} \setminus \{v\}, \quad B = N_G(v) \cap G^{-}, \quad C = G^{-} \setminus B.$$ 

Hence $A \cup \{v\}$ forms a clique and $B \cup C$ is a stable set in $G$. We can assume that $A \neq \emptyset$, for otherwise also $C = \emptyset$ and one quickly sees that $\mathcal{C}(G) \leq 3$. Let $G_v = \text{gpr}_e(G)$. Then $A$ and $C$ induce negative discrete graphs in $G_v$, $B$ induces an positive clique, and there are no edges between $B$ and $C$.

We can assume that $B \neq \emptyset$, since otherwise $G_v$ is negative connected graph, and, by Theorem 4.3, $\mathcal{C}(G) \leq \mathcal{C}(G_v) + 1 \leq (n - 1)/4 + 3 < n/2 + 4$.

(1) Let $a \in A$ be such that $\{a, c\} \notin E(G)$ for all $c \in C$. By the minimality assumption on $v$, we have $B \subseteq N_G(a)$, and therefore $a$ is isolated in $G_v$.

Notice that if $C = \emptyset$ then this case holds for all $a \in A$, and then $G_v$ consists of a positive clique and a set of isolated vertices. In this case one easily sees that $\mathcal{C}(G_v) \leq 2$, and hence $\mathcal{C}(G) \leq 3$.

Assume thus that $C \neq \emptyset$.

(2) Let $a \in A$ be such that $\{a, c\} \in E(G)$ for some $c \in C$. Then $\{a, c\} \in E(G_v)$ with $a, c \in G_v^{-}$, and if $\{a, b\} \in E(G)$ for any $b \in B$ then $\{a, b\} \notin E(G_v)$, and $\text{gdr}_{a, c}$ and $\text{gpr}_a$ can be applied in parallel to $G_v$ giving a reduction of four vertices in two steps. Also, if there exists a vertex $u \in A$ as in the above case (1) where $u$ becomes isolated in $G_v$, then $\text{gdr}_{a, c}$ and $\text{gnr}_a$ can be applied in parallel

$$\mathcal{C}(G) \leq n/2 + 4.$$
to $G_v$. Therefore we can assume that $\{a, b\} \notin E(G)$ for all $a \in A$ and $b \in B$.

By connectivity, each $x \in C$ is adjacent to at least one vertex in $A$, and no other vertices, and thus $\{a, x\} \in E(G)$ for all $x \in C$ or we have two parallel applications of the double rule by Theorem 3.2(iii). In conclusion, we have in $G$:

$$N_G(v) = B, \ N_G(a) \setminus A = C \cup \{v\}$$

for each $a \in A$, $B \cup C$ is a stable set.

In this case, $\mathcal{C}(G) \leq 4$. Indeed, the first step of such a strategy for $G$ applies rule $\text{gpr}_a$ for some $a \in A$. Then, in $\text{gpr}_a(G)$, both $A' = (A \setminus \{a\}) \cup \{v\}$ and $B$ induce negative discrete graphs, $C$ induces an positive clique, and $N_{G_v}(v) = B \cup C$.

Applying once $\text{gdr}_e$ for an edge $e = vb$ for $b \in B$ gives a disjoint union of a discrete graph and an positive clique $C$. Such a graph can be easily reduced in two steps to the empty graph.

We formulate the main open problem of this area, in three different setups.

**Problem 2.** Does there exists a constant $k$ such that $\mathcal{C}(G) \leq k$ whenever $G$ is
(a) a signed tree, (b) a negative graph, (c) any signed graph?

## 5 Graphs avoiding one type of operations

A step towards solving Problem 2 is performed by an attempt to find high complexity graphs that avoid one basic type of rule, either $\text{gnr}$, or $\text{gpr}$, or $\text{gdr}$, in all reduction strategies. The approach is clear for the double rule: since $\text{gdr}$ removes two vertices with every application, avoiding $\text{gdr}$-rules in all strategies may lengthen the reduction strategies. We investigate this problem in the following, for all three operations.

Perhaps unexpectedly, determining whether all reduction strategy of a signed graph $G$ avoid the negative rule turns out to be a difficult problem. Recall that a negative rule $\text{gnr}_v$ simply removes the negative isolated vertex $v$. The signed graphs that have no reductions using the $\text{gnr}$-rules are those that can be reduced using only $\text{gpr}$ and $\text{gdr}$, for which a string-based characterization was given in [5], but giving a graph-based characterization remains an open problem.

**Problem 3.** Give a characterization of the graphs that avoid the negative rules in all reduction strategies.

On the other hand, the case of the positive rules has a simple characterization shown in the next lemma.

**Lemma 5.1.** A signed graph $G$ has no reductions using any positive rules $\text{gpr}_v$ if and only if $G$ is negative.

**Proof.** If a signed graph $G$ has a positive vertex $v$, it can be removed only by $\text{gpr}_v$ or its sign is changed to negative by another positive rule $\text{gpr}_u$.

Let us consider now the case of $\text{gdr}$.

**Lemma 5.2.** Let $G$ be a connected signed graph with no reduction using a double rule. Then $G = G^+ \otimes G^-$, where $G^+$ is either a clique, or a disjoint union of two cliques, and $G^-$ is discrete. In this case, $\mathcal{C}(G) \leq 3$. 

9
In the above, we allow $G^- = \emptyset$, in which case $G$ is an positive clique. The conclusion can be rewritten in the form $G = (K \oplus K') \oplus S$, where $K, K'$ are positive cliques and $S$ is a negative discrete graph, where we allow that $K, K'$, and $S$ can be empty.

**Proof of the lemma.** If there is an induced path $P_3$ of three vertices in $G^+$, then by applying $\text{gpr}_v$ to the middle vertex $v$, we obtain a negative edge $e$, and so $\text{gdr}_e$ can be applied to $\text{gpr}_e(G)$; a contradiction. Obviously, if a connected graph does not have an induced path $P_3$, then it is a clique. Therefore, $G^+$ is a disjoint union of cliques, say $G^+ = G_1 \oplus \cdots \oplus G_k$.

If $G^- = \emptyset$, then $G$ is an positive clique, since it is connected. Assume thus that $G^- \neq \emptyset$. Clearly, $G^-$ must be discrete. Let $a \in G^-$. Assume that $v \in N_G(a)$ and $v, u \in G^+$ with $\{v, u\} \subseteq E(G)$. Also $u \in N_G(a)$, since $\{a, v\}$ is a negative edge in $\text{gpr}_u(G)$. This shows that if $N_G(a) \cap G_i \neq \emptyset$, then $G_i \subseteq N(a)$.

Furthermore, $G$ does not have an induced bipartite graph $K_{1,m}$ with at least three positive leaves. Indeed, if $v, w \in N_G(a)$ are not pairwise adjacent, then in $a, v, w \in \text{gpr}_w(G)^+$, and they are in the same connected component. However, for $v, u \in \text{gpr}_w(G)$ and $\text{gdr}_v$ will be applicable to $\text{gpr}_a(\text{gpr}_w(G))$; a contradiction. Consequently, any negative vertex is connected to at most two positive cliques in $G$.

Finally, suppose that there exist $a \in G^-$ and $u \in G^+$ are such that $\{a, u\} \notin E(G)$. Since $G$ is connected, there is a path from $a$ to $u$ in $G$. Choose $a$ and $u$ so that their path distance is as small as possible, and let such a shortest path be $a \rightarrow v \rightarrow \cdots \rightarrow b \rightarrow u$, where $v \in G^+$ and $b \in G^-$. By the minimality assumption on $a$ and $u$, we have $\{b, v\} \subseteq E(G)$, and thus the vertices $a, v, b, u$ induce a path $P_3$. But now $\text{gpr}_a(\text{gpr}_b(G))$ contains the negative edge $\{a, v\}$; contradiction. Hence $a$ is connected to each vertex of $G^+$, and consequently $G^+$ consists of at most two cliques. This proves the first part of the claim.

For the complexity part, note that if $G^+$ is a clique, then $\text{gpr}_u(G)$, for any $v \in G^+$, is a disjoint union of a positive clique and a discrete negative graph. Here $\ell(\text{gpr}_u(G)) \leq 2$. On the other hand, if $G^+$ is a disjoint union of two cliques and $v$ and $u$ are in different components of $G^+$, then $\text{gpr}_v$ and $\text{gpr}_u$ can be applied in parallel to $G$ and $\text{gpr}_v \text{gpr}_u(G)$ is a discrete negative graph which is reducible in one parallel step.

**Theorem 5.3.** A signed graph $G$ has no reductions using double operations if and only if the connected components of $G$ are double split graphs where the negative vertices form the stable set.

**Proof.** This is now clear from the previous lemma and by the fact that if $G$ is of the stated form, so is $\text{gpr}_v(G)$ for all positive $v \in G^+$.

A related question is to determine which graphs avoid one type of operation in at least one strategy, rather than in all strategies. For the positive rules, the answer is the same as in Lemma 5.1. In the case of the negative and double rules, the problem remains open. However, the case for the negative rules leads to a problem of independent interest. The question is formulated in Problem 4 in a more general form. Note that this invariant result holds in a stronger form for signed double occurrence strings, see [6].

**Problem 4.** Is it true that the number of negative rules ($\text{gnr}_v$ with $v \in G$) does not depend on the chosen reduction strategy for a given signed graph $G$?
6  Graphs with no parallelism in the first step

A different approach to find high complexity signed graphs is to consider graphs
for which no two operations can be applied in parallel, at least in the first step
of any reduction.

The following result is clear from Theorem 3.2 and Example 2.

**Lemma 6.1.** Let $G$ be a signed graph.

(i) $G$ has no two parallel applications of negative rules if and only if $G$ has
at most one isolated negative vertex.

(ii) $G$ has no two parallel applications of gpr-rules if and only if $G^+$ is a clique.

Moreover, if $G$ is positive, then $\mathcal{C}(G) \leq 2$.

An induced subgraph $H$ of a (signed) graph $G$ is a shadow of a vertex set (or
a subgraph) $A \subseteq G$ if for each $x \in A$ and each edge $\{v, u\}$ of $H$, $x$ is adjacent
to $v$ or $u$ or both, and each isolated vertex of $H$ is adjacent to a vertex in $A$.

**Theorem 6.2.** Let $G$ be a signed graph of at least two vertices. Then $G$ has no
parallel applications of the rules (gnr, gpr, and gdr) if and only if

(i) $G^+$ is a clique,

(ii) $G^-$ is a shadow of $G^+$, and

(iii) $G^-$ consists of a discrete graph and a complete tripartite graph.

In this case, if $G$ is negative or positive, then $\mathcal{C}(G) \leq 2$.

**Proof.** Suppose no two rules can be applied in parallel to $G$. Condition (i)
follows from Lemma 6.1 and condition (ii) obviously holds since no rule gpr$_v$
can be applied in parallel with a gnr$_u$ or with a gdr$_e$. Condition (iii) follows
from Corollary 4.2. The claim for the complexity for the cases $G = G^+$ and
$G = G^-$ is clear.

The converse claim is equally clear. □

The following simple example shows that the complexity claim of the previous
theorem does not hold if the graph is not uniformly signed.

**Example 3.** The graph in Figure 3 satisfies the conditions of Theorem 6.2.
However, its parallel complexity is three. Note that this graph admits no two
operations applied in parallel at any step of any reduction.

![Figure 3: A signed graph with no parallelism in any step of any reduction.](image-url)
A further problem is find those signed graphs that admit no two rules in parallel at any step of any reduction. With such a strong condition, our investigations so far show that only small graphs satisfy it.

**Problem 5.** Characterize the graphs for which the set of parallel reductions is identical with the set of sequential reductions.

### 7 Negative bipartite graphs of complexity one

In this section we consider the problem of characterizing those graphs $G$ that have parallel complexity one. Not surprisingly, the problem is difficult to answer in general: it is equivalent to the older open problem of characterizing the negative graphs to which a given set of double rules may be applied in parallel. We give a complete answer in several forms in this section for negative bipartite graphs. This is sufficient for the proof in the next section that negative trees have complexity at most two.

Recall that a (signed) graph $G$ is $(A,B)$-bipartite, if $A$ and $B$ are stable sets in $G$.

**Lemma 7.1.** Let $G$ be a negative $(A,B)$-bipartite graph with $A = \{a_1, \ldots, a_n\}$ and $B = \{b_1, \ldots, b_n\}$, such that $\{a_i, b_i\} \in E(G)$ and $N_G(a_i) \subseteq \{b_1, \ldots, b_i\}$ for all $1 \leq i \leq n$. Then $M = \{gdr_{a_i, b_i} | 1 \leq i \leq n\}$ can be applied in parallel to $G$. Therefore $\mathcal{C}(G) = 1$.

**Proof.** We prove the claim by induction on the order of the graphs. The result clearly holds for $n = 1$. Therefore assume that $n > 1$.

For each $i$, the graph $G_i = gdr_{a_i,b_i}(G)$ is bipartite. Indeed, $N_G(a_i) \subseteq B$ and $N_G(b_i) \subseteq A$, and hence no edges are created inside $A\{a_i\}$ or $B\{b_i\}$. Moreover, for any new edge $\{a_j, b_k\} \in E(G_i) \setminus E(G)$, we have $k \leq j$. Indeed, $b_k \in N_G(a_i)$ and $a_j \in N_G(b_i)$ and so, $k \leq i \leq j$. Consequently, $\{gdr_{a_i, b_i} | 1 \leq j \leq n, j \neq i\}$ is applicable in parallel to $G_i$, for all $1 \leq i \leq n$. Hence, by definition, the set of operations $M$ can be applied in parallel to $G$. 

A vertex $v$ of a signed graph $G$ is a leaf, if it has exactly one neighbor. Also, if $M \subseteq E(G)$ is a matching of $G$, then a cycle $\gamma$ of $G$ is said to be alternating (or $M$-alternating), if every other edge in $\gamma$ belongs to $M$ and every other edge does not.

**Lemma 7.2.** Let $M$ be a matching in the negative bipartite graph $G$ such that $gdr_M = \{gdr_e | e \in M\}$ is applicable in parallel to $G$. Then $G$ does not have $M$-alternating cycles.

**Proof.** Let $G$ be $(A,B)$-bipartite. Assume that $\gamma = e_1 f_1 \ldots e_n f_n$ is a minimum length $M$-alternating cycle in $G$, where $e_i = \{a_i, b_i\} \in M$ with $a_i \in A$, $b_i \in B$, and $f_i \notin M$. Then $\gamma$ is induced, i.e., there are no chords $\{a_i, b_j\} \in E(G)$ in $\gamma$.

If $n = 2$, then $\gamma = e_1 f_1 e_2 f_2$ is a square $C_4$, which contradicts Theorem 3.2(iii). Therefore $n \geq 3$. Then, for any $e \in M$, $gdr_e(G)$ contains an $(M \{e\})$-alternating cycle of $2n - 2$ edges, and $M \{e\}$ is a matching that is applicable in parallel to $gdr_e(G)$. Thus the claim follows inductively on the order of the graph $G$.

**Lemma 7.3.** Let $G$ be a negative bipartite graph which is not discrete. If $\mathcal{C}(G) = 1$, then $G$ has a leaf.
Proof. Let $G$ be $(A, B)$-bipartite. We can assume that $G$ is connected. Let $\mathcal{C}(G) = 1$. Therefore there exists a perfect matching $M \subseteq E(G)$ such that $\text{gdr}_M(G) = \emptyset$. Suppose that $G$ does not have any leaves. Let $\{a_1, b_1\} \in M$ with $a_1 \in A$ and $b_1 \in B$. Since $b_1$ is not a leaf, there exists an edge $\{a_2, b_1\} \in E(G) \setminus M$ and $\{a_2, b_2\} \in M$ for some $b_2 \in B$, since $\text{gdr}_M(G) = \emptyset$. Proceeding inductively we obtain an $M$-alternating cycle in $G$ which contradicts Lemma 7.2.

**Theorem 7.4.** Let $G$ be a negative $(A, B)$-bipartite graph. Then $\mathcal{C}(G) = 1$ if and only if there are orderings $A = \{a_1, \ldots, a_n\}$ and $B = \{b_1, \ldots, b_n\}$ such that $\{a_i, b_i\} \in E(G)$ and $N_G(a_i) \subseteq \{b_1, \ldots, b_i\}$, for all $1 \leq i \leq n$.

Proof. The condition was proved to be necessary in Lemma 7.1. For the reverse implication, note first that the claim is trivial $n = 1$. Let then $n > 1$, and suppose that $G$ has a perfect matching $M \subseteq E(G)$ so that $\text{gdr}_M(G) = \emptyset$.

By Lemma 7.3, $G$ has a leaf $v$. Let $u \in G$ be the unique vertex such that $e = \{v, u\} \in E(G)$. Then necessarily $e \in M$, and also $\mathcal{C}(G \setminus \{v, u\}) = 1$.

By the induction hypothesis, $G \setminus \{v, u\}$ has orderings $A' = \{a_1, \ldots, a_{n-1}\}$, $B' = \{b_1, \ldots, b_{n-1}\}$ as claimed in the theorem. Now, if $v \in A$, then let $a_0 = v$, $b_0 = u$ and if $v \in B$, then let $a_n = v$, $b_n = u$.

Clearly, these orderings satisfy the claim of the theorem.

**Corollary 7.5.** Let $G$ be a negative bipartite graph with $2n$ vertices, of parallel complexity one. Then

(i) $G$ has at least two leaves, one in both partitions;

(ii) $G$ has at most $n(n + 1)/2$ edges;

We can also characterize now the bipartite graphs with parallel complexity one in terms of alternating cycles.

**Theorem 7.6.** A negative bipartite graph $G$ has parallel complexity one if and only if $G$ has a perfect matching without alternating cycles.

Proof. Suppose $\mathcal{C}(G) = 1$, and thus that $G$ has a perfect matching $M$ such that $\text{gdr}_M(G) = \emptyset$. Now Lemma 7.2 gives the claim: there are no $M$-alternating cycles in $G$.

In converse, suppose that $G$ has a perfect matching $M$ such that $G$ has no $M$-alternating cycles. It follows that for each edge $e \in M$, we have $M \setminus \{e\} \subseteq E(\text{gdr}_e(G))$, and thus $N = M \setminus \{e\}$ is a perfect matching of $\text{gdr}_e(G)$. We show that $\text{gdr}_e(G)$ has no $N$-alternating cycles. The claim follows from this by induction on the order of the graphs.

Suppose $H = \text{gdr}_e(G)$ does have an alternating cycle $\gamma = e_1f_1e_2f_2 \ldots e_nf_n$, where $e_i = \{a_i, b_i\} \in N$ for $i = 1, 2, \ldots, n$. Let also $a_{n+1} = a_1$. Then there exists an edge $f_i = \{b_i, a_{i+1}\} \in E(H)$ such that $f_i \notin E(G)$, since $\gamma$ is not a cycle in $G$, and $N \subseteq M$. Therefore $e = \{v, u\}$ and there are edges $e, \{u, b_i\} \in E(G)$.

In $G$ there is thus the path $\{b_i, u\}e_1\{v, a_{i+1}\}$, and since $G$ does not have $M$-alternating cycles, there must be another edge $f_j = \{b_j, a_{j+1}\} \notin E(G)$. We can assume that $j < i$ and that $f_k \in E(G)$ for all $k$ with $j < k < i$. However, if $\{v, a_{j+1}\} \in E(G)$, then $\{v, a_{j+1}\}e_{j+1} \ldots e_i\{b_i, u\}e$ is an $M$-alternating cycle in $G$, and if $\{u, a_{j+1}\} \in E(G)$, then $\{u, a_{j+1}\}e_j \ldots e_i\{b_i, v\}$ is a cycle of odd length in $G$. The latter case will not occur because $G$ is bipartite. Hence $\text{gdr}_e(G)$ has no $N$-alternating cycles and this proves the claim.
From Theorem 7.6 and Theorem 7.4, we deduce

**Theorem 7.7.** A negative bipartite graph $G$ has parallel complexity one if and only if $G$ has a unique perfect matching.

**Proof.** Let $G$ be $(A, B)$-bipartite for $A = \{a_1, \ldots, a_n\}$ and $B = \{b_1, \ldots, b_n\}$ as in Theorem 7.4, and let $M$ be any perfect matching of $G$. Then $\{a_1, b_1\} \in M$, since $N_G(a_1) = \{b_1\}$. If $\{a_i, b_i\} \in M$ for all $i < j$, then also $\{a_j, b_j\} \in M$, since $N_G(a_j) \subseteq \{b_1, b_2, \ldots, b_j\}$ and the vertices $b_1, \ldots, b_{j-1}$ have already been matched. This proves the uniqueness.

In converse, if $M$ is a perfect matching of $G$ and there is an $M$-alternating cycle $\gamma = e_1f_1e_2f_2\ldots e_nf_n$, where $e_i = \{a_i, b_i\} \in M$, then also

$$N = (M \setminus \{e_1, e_2, \ldots, e_n\}) \cup \{f_1, f_2, \ldots, f_n\}$$

is a perfect matching. Hence if $G$ has a unique perfect matching $M$, then $G$ has no $M$-alternating cycles, and so $\mathcal{C}(G) = 1$ by Theorem 7.6.

For graphs that are not bipartite, the claim is not true. To see this, we can take the clique $G = K_{2n}$.

**Corollary 7.8.** A non-discrete negative tree has parallel complexity one if and only if it has a perfect matching.

### 8 Complexity of trees

#### 8.1 Negative trees

We prove in this section that any negative tree can be successfully reduced in parallel in two steps.

The following lemma is clear from the definition of the double rule.

**Lemma 8.1.** Let $e = \{a, b\}$ be an edge of $gdr_{v,u}(G)$ that is not an edge of $G$. Then $\{a, b\} \cup \{v, u\}$ induces a path $P_3$ in $G$.

Recall that an edge $e \in E(G)$ is a bridge, if the removal of $e$ disconnects $G$. An edge is a bridge if and only if it belongs to no cycle of $G$; see e.g. [21]. In the following lemma we show that bridges are inherited by the double rules.

**Lemma 8.2.** Let $e = \{a, b\}$ be a bridge in $G$. Then $e$ is a bridge of $gdr_{v,u}(G)$ for each edge $\{v, u\} \in E(G)$, where $v, u \notin \{a, b\}$.

**Proof.** Let $H = gdr_{v,u}(G)$. Since $e = \{a, b\} \in E(G)$ is a bridge and $\{v, u\} \in E(G)$, it is clear that $\{a, b\} \not\subseteq N_G(v) \cup N_G(u)$, for otherwise $\{a, b, v, u\}$ gives a cycle in $G$. Hence $e \in E(H)$. Suppose then that $e$ is not a bridge in $H$. Then there exists a path in $H$ from $a$ to $b$ that does not contain $e$. If an edge $f$ on this path does not belong to $G$, then by Lemma 8.1, the ends of $f$ are still connected by a path of length four in $G$. From this it follows that there is a connecting path of $a$ and $b$ in $G$ that does not use the edge $e$; a contradiction.

**Lemma 8.3.** For any negative graph $G$, if $M$ is a matching of $G$ consisting of bridges only, then $gdr_M$ is applicable in parallel to $G$. 

14
Proof. By Lemma 8.2, for each edge $e \in M$, we have that $M \setminus \{e\}$ is a matching of $\text{gdr}_e(G)$ also consisting of bridges only. Consequently, $\text{gdr}_M$ can be applied sequentially and in any order on the edges of $M$. Thus $\text{gdr}_M$ is applicable in parallel to $G$. 

Let $M$ be a matching of a graph $G$. We say that a path $P = e_1e_2\ldots e_m$ is $M$-alternating, if $e_{2i-1} \notin M$ and $e_{2i} \in M$ for all $i \geq 1$. Such a path $P$ is $M$-augmented, if the ends of the path are not saturated by $M$, i.e., if the path starts from $a$ and ends in $b$, then $a$ and $b$ are not endpoints of any edge $e \in M$.

**Theorem 8.4.** Let $M$ be a maximum matching of a negative graph consisting only of bridges of $G$. Then $\text{gdr}_M(G)$ is discrete.

**Proof.** First of all, $\text{gdr}_M(G)$ does not have any edges of $G$, since $M$ is maximum, and the ends of the edges in $M$ are removed from $G$. Let $M = \{e_1, e_2, \ldots, e_m\}$, and denote $G_0 = G$ and $G_i = \text{gdr}_{e_i}(G_{i-1})$ for all $i = 1, 2, \ldots, m$. Suppose that $(v, u)$ is an edge in $G_i$ but not in $G_{i-1}$. Hence $(v, u)$ is created by applying $\text{gdr}_{e_i}$ to $G_{i-1}$ for $e_i = \{a, b\} \in M$, and so there is an $M$-alternating path in $G_{i-1}$ from $v$ to $u$. Indeed, this path is either $(v, a, b, u)$ or $(v, b, a, u)$. Inductively, we obtain that there is an $M$-alternating path in $G_i$ from $v$ to $u$ for each $j < i$, and therefore also in $G$. We conclude that one of the ends, $v$ or $u$, must be saturated by $M$ by Berge’s criterion for maximum matchings; see [21]. The claim follows from this by applying the above to the final result $G_m = \text{gdr}_M(G)$. 

**Corollary 8.5.** Any negative tree can be reduced in one step to a discrete graph and so, its complexity is at most two.

### 8.2 Positive trees

In this section we prove that any positive tree can be reduced in one step to a negative graph that may not be a tree anymore. However, we are able to prove that these graphs have a special form and in particular, they have complexity at most two.

The following purely graph theoretic result is interesting on its own.

**Lemma 8.6.** Every tree $T$ has a stable set $I$ such that the vertices $v \notin I$ are adjacent to an odd number of elements of $I$.

**Proof.** The proof goes by induction on the number of vertices in the tree. The claim is clear for trees of size at most 4. Suppose the claim holds for all trees of at most $n-1$ vertices. Let $T$ be a tree of $n$ vertices. We divide the proof to several cases that exhaust all possibilities.

We say that a subset $I \subseteq T$ is good if $I$ is stable and each $v \in G \setminus I$ is adjacent to an odd number of elements of $I$.

(1) Assume $T$ has a vertex $v$ adjacent to at least two leaves, say $x$ and $y$.

Consider the subtree $T' = T \setminus \{x, y\}$ which has, by the induction hypothesis, a good subset $I'$. If $v \in I'$, then clearly $I = I'$ is a good subset of $T$, since $x$ and $y$ are adjacent in $T$ to only one element of $I$, namely $v$. On the other hand, if $v \notin I'$, then $I = I' \cup \{x, y\}$ is a good subset of $T$, since it is stable and since $v$ is adjacent to an even number (two) new neighbors, not in $I'$.

There remains the case where each vertex of $T$ is adjacent to exactly one leaf or no leaves at all. A path $v_1, v_2, \ldots, v_k$ is called a branch, if $v_k$ is a leaf
Let \( T \) be applied in parallel to \( T \) only vertex of \( \text{gpr}(T) \) gives a good subset \( I' \) of \( T' \). If \( v_{k-3} \in I' \) then \( I = I' \cup \{ v_k \} \) is good for \( T \), and if \( v_{k-3} \notin I' \) then \( I = I' \cup \{ v_{k-1} \} \) is good for \( T \).

(3) Hence we can assume that the branches are all of length \( k \leq 3 \). In this case there exists a vertex \( v \in T \) such that \( v \) starts at least two branches and the degree of \( v \) is \( |N_T(v)| \geq 3 \).

(3.1) Suppose first that there are different branches \( v, x_1, x_2 \) and \( v, y_1, y_2 \) both of length two. By the induction hypothesis, the subtree \( T' = T - \{ x_1, x_2 \} \) has a good subset \( I' \). Now \( v \notin I' \), since otherwise \( y_1 \notin I' \) and so also \( y_2 \notin I' \) contradicting the fact that \( y_2 \) is adjacent to an element of \( I' \). Therefore, \( I = I' \cup \{ x_2 \} \) is good for \( T \).

(3.2) Finally suppose that there are different branches \( v, x_1 \) and \( v, y_1, y_2 \). By the induction hypothesis \( T' = T - \{ x_1 \} \) has a good subset \( I' \). As in the previous case \( v \notin I' \). In this case, either \( y_1 \in I' \) or \( y_2 \in I' \). If \( y_1 \in I' \), then \( y_2 \notin I' \), and \( I = (I' \setminus \{ y_1 \}) \cup \{ x_1, y_2 \} \) is good for \( T \). On the other hand, if \( y_2 \in I' \) and so \( y_1 \notin I' \), then \( I = (I' \setminus \{ y_2 \}) \cup \{ x_1, y_1 \} \) is good for \( T \). This proves the claim.

We say that a stable subset \( I \subseteq T \) is a signer subset for the positive tree \( T \) if \( \text{gpr}_I(T) \) is a negative graph. Note that \( \text{gpr}_I(T) \) need not be a tree.

**Lemma 8.7.** Let \( T \) be a positive tree. Then it has a signer subset \( I \subseteq T \). Moreover, if \( I \) is a signer subset, then \( \text{gpr}_I(T) \) can be reduced to a discrete graph in one parallel step.

Before proving Lemma 8.7, we need to state two auxiliary results about all-negative graphs.

**Lemma 8.8.** Let \( G \) be an all-negative graph, \( K \) a clique of \( G \) and \( u_0 \in K \) the only vertex of \( K \) adjacent to vertices in \( G \setminus K \).

(i) For any edge \( uv \) of \( K \), \( u \notin \{ v, w \} \), \( \text{gdr}_{v,w}(G) = G \setminus \{ u, v \} \);

(ii) For any edge \( xy \) in \( G \setminus K \), \( K \) is a clique in \( \text{gdr}_{x,y}(G) \) and \( u_0 \) is the only vertex of \( K \) adjacent to vertices in \( \text{gdr}_{x,y}(G) \setminus K \);

(iii) For any edge \( uvw \) of \( K \), \( K \setminus \{ u_0, v \} \) is a clique in \( G' = \text{gdr}_{u_0,v}(G) \) and for all \( u \in K \setminus \{ u_0, v \}, N_{G'}(u) \cup \{ u \} = N_G(u_0) \setminus \{ v \} \);

(iv) For any edge \( u_0x \) of \( G \), with \( x \notin K \), \( K \setminus \{ u_0 \} \) is a clique in \( G'' = \text{gdr}_{u_0,x}(G) \) and for all \( u \in K \setminus \{ u_0 \}, N_{G''}(u) \cup \{ u \} = N_G(x) \cup (K \setminus \{ u_0 \}) \).

**Lemma 8.9.** Let \( G \) be an all-negative graph and \( u, v \) twins in \( G \). Then

(i) \( u \) and \( v \) are twins in every \( \text{gdr}_{x,y}(G) \), whenever \( u, v \notin \{ x, y \} \), and

(ii) \( \text{gdr}_{u,v}(G) = G \setminus \{ u, v \} \).

**Proof of Lemma 8.7.** If \( I \) is a stable subset of \( T \), then \( \text{gpr}_I = \{ \text{gpr}_v \mid v \in I \} \) can be applied in parallel to \( T \), and so \( G = \text{gpr}_I(T) \) is well defined. By Lemma 8.6, there exists a stable subset \( I \subseteq T \) such that each \( v \notin I \) has an odd number of
neighbors in $I$. Therefore each sign $v \in T$ is changed to negative in $\text{gpr}_f(T)$, and hence $G = \text{gpr}_f(T)$ is a negative graph. Hence $I$ is a signer subset of $T$.

For the second part of the theorem, let $I = \{v_1, \ldots, v_m\}$ be a signer subset of $T$. Since $T$ is a tree, the neighborhood $K_i = N_G(v_i)$ of each $v_i \in I$ is a stable set, and thus it induces a clique in $G$. Moreover, in order to avoid cycles in $T$, we must have that $|K_i \cap K_j| \leq 1$ for all different $i$ and $j$. In case $K_i \cap K_j = \emptyset$, there is at most one pair of vertices $v \in K_i$, $u \in K_j$ with $\{v, u\} \in E(G)$, since in this case also $\{v, u\} \in E(T)$. In all such cases, we add to our list of cliques the clique $\{u, v\}$. Let $K_1, \ldots, K_n$, $n \geq m$ be our final list of cliques, where there is no edge between vertices in distinct cliques.

Observe now that there are no distinct vertices $u_1, \ldots, u_k$, $k \geq 3$, such that $u_l \in K_i \cap K_{i+l}$, $1 \leq l \leq k-1$, and $u_k \in K_i \cap K_{i+1}$, for some $1 \leq i_1, \ldots, i_k \leq n$. We say that the cliques $K_1, \ldots, K_m$ induce a forest-like ‘meta-structure’ in the following formal sense: (i) the cliques $K_1, \ldots, K_n$ are intersecting on at most singletons, and (ii) all simple cycles of $G$ are simple cycles of some clique $K_j$, $1 \leq j \leq n$.

We prove the claim of the lemma by induction on the number of vertices in the graph, for all graphs that are forest-like, as defined above.

For the inductive step, note that it follows directly from the definition of forest-like that there is a clique, say $K_1$, that is a ‘leaf’ in our meta-structure: there is at most one vertex $v_0 \in K_1$ such that $v_0 \in K_i$, for some $i \neq 1$. If $K_1$ has no such vertex $u_0$ (i.e., $K_1$ is ‘isolated’), then applying $\text{gdr}$ on any maximal matching for $K_1$ will reduce $K_1$ to a discrete graph, and it may be applied in parallel with any other $\text{gdr}$ operations applied on edges of $G \setminus K_1$. Assume now that $K_1$ is not ‘isolated’ in $G$.

If $|K_1| = 2$, say $K_1 = \{u_0, v\}$, then observe that $G \setminus K_1$ has a forest-like ‘meta-structure’ and so, by induction hypothesis, there is a matching $M'$ of $G \setminus K_1$ such that $\text{gdr}_{M'}(G \setminus K_1)$ is a discrete graph. We claim that $\text{gdr}$ is applicable in parallel on $M = M' \cup \{u_0v\}$ to $G$ and $\text{gdr}_{M}(G)$ is discrete. To this aim, we must show that $\text{gdr}$ operations may be applied in any sequential order on edges in $M$. Note first that based on Lemma 8.8, throughout applying $\text{gdr}$ operations on edges in $M'$, $K_1$ remains a clique, with $u_0$ its only vertex adjacent to vertices outside $K_1$. Consequently, $\text{gdr}$ may be applied on $u_0v$ at any point, resulting only in the removal of vertices $u_0$ and $v$. Then $\text{gdr}$ may be applied on the remaining edges of $M'$ in an arbitrary order, proving the claim.

If $|K_1| > 2$, then let $u, v \in K_1 \setminus \{u_0\}$. Observe that $G \setminus \{u, v\}$ has a forest-like ‘meta-structure’ and so, by induction hypothesis, there is a matching $M''$ of $G \setminus \{u, v\}$ such that $\text{gdr}_{M''}(G \setminus \{u, v\})$ is a discrete graph. We claim that $\text{gdr}$ is applicable in parallel on $M = M'' \cup \{uv\}$ to $G$ and $\text{gdr}_{M}(G)$ is discrete. It follows from our choice of $u, v$ that $N_G(u) \cup \{u\} = N_G(v) \cup \{v\}$. Consequently, based on Lemmas 8.8 and 8.9, throughout applying $\text{gdr}$ operations on any edges of $M''$, the edge $uv$ is preserved and their closed neighborhoods remain the same. Thus, $\text{gdr}_{u,v}$ is applicable at any point, resulting only in the removal of $u, v$. Then $\text{gdr}$ may be applied on the remaining edges of $M''$ in an arbitrary order, proving the claim. 

9 Discussion

Although the parallel complexity of negative and positive trees is bounded, the question of the bound for that of the signed trees still remains open. The reason is that by local complementing, a signed tree may be transformed into an arbitrary graph. The following table lists the smallest signed trees with respect to the complexity from one to five. The result is based on a full search that explores all possible signed trees of less than 18 vertices. Here, \( n \) is the size of the tree, \( C \) is the parallel complexity corresponding to the tree. See examples and proofs in [11].

<table>
<thead>
<tr>
<th>( C )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n )</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>6</td>
<td>12</td>
</tr>
</tbody>
</table>

Since there is no known efficient way to decide the maximum set of negative edges that can be reduced in parallel for a given set of edges \( S \) in a given graph \( G \), a bound on the parallel complexity for negative graphs is not concluded. Here, we recall the conjecture in [9], that negative graphs only have parallel complexity up to three. Indeed, it is possible to find negative graphs of parallel complexity three, while negative graphs of complexity higher than that is not found through an automated search [23]. Interestingly, the negative graphs found of parallel complexity three are tripartite graphs. [11].

Settling the complexity of positive graphs seems to be a difficult problem. The main reason for this is that this class of graphs is not closed under \( gpr \). Some positive tripartite graphs of parallel complexity four have been reported in [11], the highest complexity currently known for this class.

Graphs with parallel complexity up to six were reported in [11]. Proving the complexity of those graphs is based essentially on a computer-based exhaustive search implemented in [23], see [12] for a description of the algorithm.

Acknowledgment The authors gratefully acknowledge support by Academy of Finland (TH – project 39802, CL – project 203667, IP – project 108421).

References


