Combating Collusion Rings is Hard but Possible

Niclas Boehmer[1], Robert Bredereck[2], and André Nichterlein[1]

[1] Technische Universität Berlin, Faculty IV, Algorithmics and Computational Complexity, Berlin, Germany
{niclas.boehmer,andre.nichterlein}@tu-berlin.de
[2] Humboldt-Universität zu Berlin, Institut für Informatik, Algorithm Engineering, Berlin, Germany
robert.bredereck@hu-berlin.de

December 15, 2021

Abstract

A recent report of Littmann [Commun. ACM '21] outlines the existence and the fatal impact of collusion rings in academic peer reviewing. We introduce and analyze the problem CYCLE-FREE REVIEWING that aims at finding a review assignment without the following kind of collusion ring: A sequence of reviewers each reviewing a paper authored by the next reviewer in the sequence (with the last reviewer reviewing a paper of the first), thus creating a *review cycle* where each reviewer gives favorable reviews. As a result, all papers in that cycle have a high chance of acceptance independent of their respective scientific merit.

We observe that review assignments computed using a standard Linear Programming approach typically admit many short review cycles. On the negative side, we show that CYCLE-FREE REVIEWING is NP-hard in various restricted cases (i.e., when every author is qualified to review all papers and one wants to prevent that authors review each other's or their own papers or when every author has only one paper and is only qualified to review few papers). On the positive side, among others, we show that, in some realistic settings, an assignment without any review cycles of small length always exists. This result also gives rise to an efficient heuristic for computing (weighted) cycle-free review assignments, which we show to be of excellent quality in practice.

1 Introduction

As recently pointed out by Littman [2021], the integrity and legitimacy of scientific conference publications (particularly important in the context of computer science) is threatened by so-called "collusion rings", which are sets of authors that unethically review and support each other while breaking anonymity and hiding conflicts of interest. Despite the fact that details are usually not disclosed for various reasons, it is inevitable that the process of assigning papers to reviewers is the key point to engineer technical barriers against such incidents. Whereas assignments at very small venues could be performed manually, support by (semi-)automatic systems becomes necessary already for medium-size conferences. Today computational support for finding review assignments is well-established and has improved the quality of the reviewing and paper assignment process in many ways (see the surveys of Shah [2021] and Price and Flach [2017] for details). Still there is huge potential for improving processes and further computational support is urgently requested [Price and Flach, 2017, Shah, 2021].

When aiming to prevent collusion rings, one of the most basic properties one can request from a review assignment is that the assignment does not contain any *review cycle* of length z, that is, a sequence of z agents each reviewing a paper authored by the next agent in the sequence (with the last agent reviewing a paper authored by the first). This property is of high practical relevance: For example, in the AAAI'21 review assignment the non-existence of review cycles of length at most $z = 2$

was a soft constraint [Leyton-Brown and Mausam, 2021]. Yet, there is a lack of systematic studies concerning the computation of such assignments. Motivated by this, we propose and analyze CYCLE-FREE REVIEWING, the problem of computing an assignment of papers to agents that is free of review cycles of length at most z, both from a theoretical and practical perspective.

1.1 Related Work

The literature is rich in the general context of peer reviewing (see, e. g., the works of Garg et al. [2010], Goldsmith and Sloan [2007], Kobren et al. [2019], Lian et al. [2018], Long et al. [2013], Stelmakh et al. [2021], Taylor [2008] on computational aspects of finding a "good" review assignment, and the survey of Shah [2021]). Closest to our work are Barrot et al. [2020] and Guo et al. [2018]. In the context of product reviewing, among others, Barrot et al. [2020] propose and analyze a restricted case which translates to our setting as follows: Given a set of single-author papers and a set of agents each writing a single paper and each having some conflicts of interest over papers, find a review assignment of papers to agents, where each agent serves as a reviewer providing one review and each paper must receive one review. They show that in this setting finding an assignment without review cycles of length at most z corresponds to finding a 2-factor without cycles of length at most z, which is known to be NP-hard for $z \geq 5$ but polynomial-time solvable for $z \leq 3$ [Hell et al., 1988]. Closer to our setting is that of Guo et al. [2018], who also consider the computation of cycle-free review assignments. They propose two simple heuristics and conduct experiments measuring the quality of their heuristics and the number of review cycles in a weight-maximizing solution on two instances, mostly focusing on the influence of the number of reviews per paper and per reviewer.

1.2 Outline and Contributions

Our contribution is threefold. First, in Section 3, we show the intractability of CYCLE-FREE REVIEWING in various restricted settings: We show NP-hardness even when just forbidding review cycles of length at most two in "sparse" and "dense" settings (e.g., if each reviewer can review only "few" or can review "almost all" papers, see Theorems 1 to 3). Furthermore, solving a question left open by Barrot et al. [2020], we show NP-hardness if each agent writes just one single-author paper and can review only few papers (Theorem 4).

Second, in Section 4, we develop greedy heuristics. In contrast to Guo et al. [2018] we provide a theoretical analysis for the heuristics. In particular, we prove that, if the considered instance satisfies certain near-realistic conditions (such as that each paper has few authors and that for each paper there are many possible reviewers), then these heuristics are guaranteed to output a z-cycle-free review assignments in polynomial time.

Third, in Section 5, we present and discuss the results of our experiments. Our core results are:

1. Existing linear-programming-based methods for computing maximum-weight review assignments (as often used in practice) produce assignments where a high fraction (20% or more) of agents and papers belong to some review cycles of length two.

2. For $z \in \{2, 3, 4\}$ maximum-weight z-cycle-free assignments computed by one of our heuristics (see Section 4) or computed via Integer Linear Programming are almost as good as the maximum-weight review assignments with cycles (solution quality loss less than 4% resp. 1%).

3. Somewhat surprisingly, we show that adding additional reviewers that are authors of some papers to the reviewer pool increases the number of papers that belong to review cycles in maximum-weight (non cycle-free) assignments.

Table 1: Notation overview

Variable	Explanation
$V = A \uplus P$	vertex set consisting of agents A and papers P with $n_A = \|A\|$ and $n_P = \|P\|$
E_A	$(a,p) \in E_A \subseteq A \times P$ shows a can review p
E_P	$(p,a) \in E_P \subseteq P \times A$ shows a authors p
$N^-(v,E)$	in-neighbors of $v \in V$ wrt. $E \subseteq \binom{V}{2}$, i. e., $N^-(v,E) := \{u \in V \mid (u,v) \in E\}$
$N^+(v,E)$	out-neighbors of $v \in V$ wrt. $E \subseteq \binom{V}{2}$, i. e., $N^+(v,E) := \{u \in V \mid (v,u) \in E\}$
Δ_U^-, Δ_U^+	maximum in- and out-degree in U resp., e. g., $\Delta_U^- := \max_{u \in U} \|N^-(u, E_A \uplus E_P)\|$
δ_U^-, δ_U^+	minimum in- and out-degree in U resp., e. g., $\delta_U^+ := \min_{u \in U} \|N^+(u, E_A \uplus E_P)\|$
Δ_A^-, δ_A^-	maximum resp. minimum number of papers per author
Δ_P^+, δ_P^+	maximum resp. minimum number of authors per paper
Δ_A^+, δ_A^+	maximum resp. minimum number of papers any author is qualified to review
Δ_P^-, δ_P^-	maximum resp. minimum number of potential reviewers for any paper

2 Preliminaries

For $n \in \mathbb{N}$, we set $[n] := \{1, \dots, n\}$. In an instance of CYCLE-FREE REVIEWING, we are given a set P of papers and set A of agents, where each paper $p \in P$ is authored by a subset $\text{aut}(p) \subseteq A$ of agents. Moreover, we are given for each agent $a \in A$ a subset $\text{rev}(a) \subseteq P$ of papers the agent is *qualified to review*[1]. We capture this information in a bipartite graph $(A \uplus P, E_A \uplus E_P)$ with $E_A = \{(a,p) \mid a \in A, p \in \text{rev}(a)\}$ and $E_P = \{(p,a) \mid p \in P, a \in \text{aut}(p)\}$ (see also Table 1 for an overview). A *(peer) review assignment* $E' \subseteq E_A$ is a subset of edges from agents to papers, where we say that a *reviews* p in E' if $(a,p) \in E'$. Given a review assignment $E' \subseteq E_A$, for an agent $a \in A$, let $N^+(a,E') = \{p \in P \mid (a,p) \in E'\}$ be the subset of papers agent a reviews in E' and, for a paper $p \in P$, let $N^-(p,E') = \{a \in A \mid (a,p) \in E'\}$ be the subset of agents that review p in E'. For $c, d \in \mathbb{N}$ a review assignment $E' \subseteq E_A$ is called *c-d-valid* if each agent reviews at most c papers and each paper is reviewed by d agents, that is, $|N^+(a,E')| \leq c$ for all $a \in A$ and $|N^-(p,E')| = d$ for all $p \in P$. In a review assignment $E' \subseteq E_A$, we say that papers $p_1, \dots, p_z$ and agents $a_1, \dots, a_z$ form a *review cycle (of length z)* if a_i is an author of p_i $\big((p_i, a_i) \in E_P\big)$ for all $i \in [z]$, a_i reviews p_{i+1} in E' $\big((a_i, p_{i+1}) \in E'\big)$ for $i \in [z-1]$ and a_z reviews p_1 in E' $\big((a_z, p_1) \in E'\big)$. Notably, a review cycle of length z in E' corresponds to a directed cycle of length $2z$ in $(A \uplus P, E' \uplus E_P)$ and a review cycle of length one corresponds to an author reviewing one of its own papers. We say that a review assignment E' is z-cycle free if there is no review cycle of length $i \in [z]$ in E'.

Using this notation, we define our central problem and refer to Table 1 for further necessary variable definitions:

[WEIGHTED] CYCLE-FREE REVIEWING
Input: A directed bipartite graph $(A \uplus P, E_A \uplus E_P)$ and non-negative integers c_{reviewer}, d_{paper}, and z [and a weight function $w : E_A \mapsto \mathbb{Z}$ and an integer W].
Question: Is there a c_{reviewer}-d_{paper}-valid and z-cycle-free review assignment $E' \subseteq E_A$ [of weight at least W, i.e., $\sum_{e \in E'} w(e) \geq W$]?

3 NP-Hardness in Various Restricted Cases

From the work of Barrot et al. [2020, Theorem 4.12] it follows that CYCLE-FREE REVIEWING is NP-hard in the single-author-single-paper setting ($\Delta_A^- = \Delta_P^+ = 1$) even if $c_{\text{reviewer}} = d_{\text{paper}} = 1$ and

[1] Being "qualified to review" can encode that the agent is capable of reviewing the paper or that the agent does not have a conflict of interest with one of the co-authors or both.

$z = 2$. However, as in reality instances of CYCLE-FREE REVIEWING are hardly arbitrary but have a quite strong structure, in this section we prove that the NP-hardness of CYCLE-FREE REVIEWING upholds even if the given instance fulfills further quite restrictive conditions, e.g., each agent is qualified to review all papers or our problem specific parameters ($\Delta_A^-, \Delta_P^+, \Delta_A^+, \Delta_P^-, c_{\text{reviewer}}, d_{\text{paper}}, z$) are small constants.

3.1 Sparse Review Graph and Small Weights

We start by considering the case where all our parameters are small. Specifically, we show the NP-hardness of CYCLE-FREE REVIEWING for arbitrarily $z \geq 2$ even if each paper is only authored by at most two agents, each agent authors at most two papers, each agent is only qualified to review at most three papers, and for each paper only at most three agents are qualified to review it (see Table 1 for definitions).

Theorem 1. *For any $z \geq 2$, CYCLE-FREE REVIEWING is NP-hard, even if $\Delta_A^+ = \Delta_P^- = 3$, $\Delta_A^- = \Delta_P^+ = 2$, $n_A = n_P$, and $c_{\mathit{reviewer}} = d_{\mathit{paper}} = 1$. The hardness results still hold if agents are not allowed to review papers of co-authors.*

Proof. We reduce from an NP-hard variant of SATISFIABILITY where each clause consists of exactly three literals and each variable occurs positive in at most two clauses and negative in at most two clauses [Berman et al., 2003].

Construction. Given an instance of SATISFIABILITY consisting of a set X of variables and a set C of clauses, we set $d_{\text{paper}} = c_{\text{reviewer}} = 1$ and z to some integer greater than one. We construct the set A of agents and the set P of papers as follows. For each variable $x \in X$, we introduce three agents a_x, $a_{\bar{x}}$, and b_x and three papers p_x, $p_{\bar{x}}$, and q_x (q_x has no author and can be considered as a dummy paper). Agents a_x and b_x are qualified to review p_x, agents $a_{\bar{x}}$ and b_x are qualified to review $p_{\bar{x}}$ and agents a_x and $a_{\bar{x}}$ are qualified to review q_x. Intuitively, either does a_x review p_x (which corresponds to setting x to false) or $a_{\bar{x}}$ review $p_{\bar{x}}$ (which corresponds to setting x to true).

For each clause $c = \ell_1 \vee \ell_2 \vee \ell_3$, we introduce three agents a_c^1, a_c^2, and a_c^3 and three papers p_c^1, p_c^2, and p_c^3 where a_c^i is qualified to review p_c^i for $i \in [3]$. Moreover, we introduce two dummy agents that are both qualified to review p_c^1, p_c^2, and p_c^3 and two dummy papers who a_c^1, a_c^2, and a_c^3 are all qualified to review. Notably, for one $i \in [3]$, a_c^i needs to review p_c^i (which corresponds to c being fulfilled because of ℓ_i).

Concerning the authors of each paper, for each clause $c = \ell_1 \vee \ell_2 \vee \ell_3$ and $i \in [3]$, a_c^i is an author of p_{ℓ_i} and a_{ℓ_i} is an author of p_c^i.

It is easy to see that each agent is only qualified to review at most three papers and that for each paper only at most three agents are qualified to review it. Moreover, as each literal only appears in at most two clauses, every paper has at most two authors and each agent authors at most two papers. Moreover, note that $|A| = |P|$, implying that each agent has to review *exactly* one paper.

(⇒) Let Z be the set of variables that are set to true in a satisfying assignment of the given SATISFIABILITY instance. Then, for $x \in Z$, we assign b_x to p_x, a_x to q_x, and $a_{\bar{x}}$ to $p_{\bar{x}}$, while for $x \notin Z$, we assign b_x to $p_{\bar{x}}$, $a_{\bar{x}}$ to q_x, and a_x to p_x. For a clause $c = \ell_1 \vee \ell_2 \vee \ell_3$, let ℓ_{i^*} with $i^* \in [3]$ be a literal from c that is set to true by the given assignment (such a literal exists because the given assignment is satisfying). Then, we set $a_c^{i^*}$ to review $p_c^{i^*}$. The two dummy agents from this clause are assigned arbitrarily to p_c^i for $i \in [3] \setminus \{i^*\}$ and the agents a_c^i for $i \in [3] \setminus \{i^*\}$ are assigned arbitrarily to the two dummy papers. To show that the constructed assignment does not contain a review cycle (of arbitrary length) note that only papers that have an author are papers p_ℓ for some literal ℓ (which are authored by a_c^i for some $c \in C$ and $i \in [3]$ where ℓ appears in c as the ith literal) and papers p_c^i for some $c = \ell_1 \vee \ell_2 \vee \ell_3 \in C$ and $i \in [3]$ (which are authored by a_{ℓ_i}). Thus, every review cycle of length at least two needs to contain an agent

a_c^i for some $c \in C$ and $i \in [3]$ and a_ℓ, where ℓ appears in c as the ith literal, and a_c^i reviews p_c^i and a_ℓ reviews p_ℓ. For a_c^i to review p_c^i it needs to hold that the given assignment satisfies ℓ. However, by our construction of the review assignment, a_ℓ reviewing p_ℓ implies that $\bar{\ell}$ is satisfied. Thus, no review cycle exists.

($\Leftarrow$) Assume we are given a 1-1-valid z-cycle-free review assignment. Let $Y := \{x \in X \mid a_{\bar{x}} \text{ reviews } p_{\bar{x}}\}$. We claim that the assignment α which sets all variables in Y to true and all variables in $X \setminus Y$ to false satisfies the given formula. Assume for the sake of contradiction that there exists a clause $c = \ell_1 \vee \ell_2 \vee \ell_3 \in C$ which is not satisfied by α. As the given assignment is 1-1-valid and we have the same number of agents and papers in the constructed instance, there is a $i^* \in [3]$ such that $a_c^{i^*}$ reviews $p_c^{i^*}$. Note that by the same reasoning, for each $x \in X$, either does a_x review p_x or $a_{\bar{x}}$ review $p_{\bar{x}}$. Thus, if a literal ℓ is not satisfied by α, then a_ℓ reviews p_ℓ. As ℓ_{i^*} is not satisfied by α , $a_{\ell_{i^*}}$ reviews $p_{\ell_{i^*}}$. Thus, a_c^i and $a_{\ell_{i^*}}$ form a review cycle of length two, as a_c^i reviews p_c^i, which is authored by $a_{\ell_i^*}$, and $a_{\ell_i^*}$ reviews $p_{\ell_i^*}$, which is authored by a_c^i, a contradiction. □

The above reduction crucially relies on the "sparsity" of the qualifications, i.e., that each agent is qualified to review between two and three papers and that for each paper only two or three agents are qualified to review it. Motivated by the observation that, in practice, reviewers are typically qualified to review more than just two or three papers and that for each paper there typically exists more than just two or three qualified reviewers, it is a natural question whether our above hardness result still extends to this case. We answer this question affirmative by proving hardness for arbitrary δ_A^+ and δ_P^-, i.e., for the case where each agent is qualified to review at least δ_A^+ papers and for each paper there exist at least δ_P^- agents that are qualified to review to:

Proposition 1. *For any $z \geq 2$, $\delta_P^- \geq 2 \leq \delta_A^+$,* CYCLE-FREE REVIEWING *is NP-hard, even if $\Delta_A^- = \Delta_P^+ = 2$, $n_A = n_P$, and $c_{\text{reviewer}} = d_{\text{paper}} = 1$.*

Proof. Let $\delta := \max(\delta_A^+, \delta_P^-)$. We reduce from the restricted NP-hard variant of CYCLE-FREE REVIEWING considered in Theorem 1. Given an instance $\mathcal{I} = ((A \cup P, E_A \cup E_P), c_{\text{reviewer}} = 1, d_{\text{paper}} = 1, z = 2)$ of CYCLE-FREE REVIEWING with $\Delta_A^- = \Delta_P^+ = 2$, we modify the instance $\mathcal{I}$ by introducing two sets A' and A'' of δ agents each and two sets P' and P'' of δ papers each. All agents from A' are qualified to review all papers from P' and from P. In addition to being qualified to review some papers from P (as captured in E_A), all agents from A are qualified to review all papers from P''. Moreover, all agents from A'' are qualified to review all papers from P''. Thereby, all agents are qualified to review at least δ papers and for each paper at least δ agents are qualified to review it. Notably, we still have $|A| = |P|$. Thus, as agents from A'' are only qualified to review papers from P'' and $|A''| = |P''|$, all papers from P'' need to be reviewed by agents from A'' (which is always possible to do in without creating a review cycle as no paper from A'' has an author). Similarly, as papers from P' can only be reviewed by agents from A' and $|A'| = |P'|$, all agents from A' need to review papers from P' (which is always possible to do in without creating a review cycle as no paper from A' has an author). Thus, all agents from A need to review papers from P from which the correctness of the reduction directly follows.

Lastly, note that we did not modify the set of authors for any paper from P and did not add papers with an author. Thus, it still holds in the modified instance that each agent authors at most two papers and each paper has at most two authors ($\Delta_A^- = \Delta_P^+ = 2$). □

While we prove hardness for arbitrary δ_A^+ and δ_P^-, in our construction from Proposition 1, there are always agents that are not qualified to review "many" papers (around $\frac{2}{3}$) and always papers that cannot be reviewed by "many" agents (around $\frac{2}{3}$). Thus, interpreting a qualification as the absence of a conflict of interest, for our NP-hardness agents need to have many conflicts. In Section 4, we prove that this does not happen by accident, as if the number of conflicts per agent/paper (and Δ_A^-, Δ_P^+, c_{reviewer}, and d_{paper}) are "small", then CYCLE-FREE REVIEWING always admits a solution.

In WEIGHTED CYCLE-FREE REVIEWING it is possible to encode the "qualifications" of agents into weights: If we modify the reduction from above and give an agent-reviewer pair weight one if the agent is qualified to review the paper and weight zero otherwise, we get that WEIGHTED CYCLE-FREE REVIEWING is NP-hard even if each agent is qualified to review all papers and we have few non-zero weights.

Corollary 1. *For any $z \geq 2$,* WEIGHTED CYCLE-FREE PEER REVIEWING *is NP-hard, even if each agent is qualified to review all papers, each agent gives only at most three papers a non-zero weight, for each paper at most three agents give it a non-zero weight, $\Delta_P^+ \leq 2 \geq \Delta_A^-$, $n_A = n_P$, and $c_{\mathit{reviewer}} = d_{\mathit{paper}} = 1$.*

3.2 No Conflicts of Interest

We now extend the hardness from Corollary 1 for the case where each agent is qualified to review all papers (no conflicts) to the unweighted case. However, our new reduction relies on the existence of papers with many authors and agents authoring many papers.

To show that CYCLE-FREE REVIEWING is NP-hard even if each agent is qualified to review all papers, $n_A = n_P$, $c_{\text{reviewer}} = d_{\text{paper}} = 1$, and $z = 2$ (Theorem 2), we reduce from MULTICOLORED INDEPENDENT SET where we are given a graph G with vertices partitioned into k sets $V^1, \ldots, V^k$ (to which we refer as color classes) and the question is whether there exists a subset of k vertices, containing one vertex from each class, that are pairwise non-adjacent. We denote as $n := |V^1|$ the number of vertices in the first color class and assume without loss of generality that $n > k$ and that $|V^c| := n + c - 1$ for $c \in [k]$ (note that we can do so because we can always add vertices that are connected to all other vertices and put them into one of the color classes).

Construction. Given an instance $\mathcal{I}$ of MULTICOLORED INDEPENDENT SET $G = (V = (V^1, \ldots, V^k), E)$, we construct an instance $\mathcal{I}'$ of CYCLE-FREE REVIEWING as follows. For each color $c \in [k]$, we add a *special agent* a_*^c and a *special paper* p_*^c. Moreover, for each vertex $v \in V^c$, we add a *vertex agent* a_v^c and a *vertex paper* p_v^c. Further, we add $n + c - 2$ *dummy agents* $\tilde{a}_1^c, \ldots, \tilde{a}_{n+c-2}^c$ and $n + c - 2$ *dummy papers* $\tilde{p}_1^c, \ldots, \tilde{p}_{n+c-2}^c$. Lastly, we insert an agent a^* and a paper p^*.

The paper p^* is authored by all vertex agents and dummy agents. For color $c \in [k]$, p_*^c is authored by all vertex und dummy agents from colors $c' \neq c \in [k]$ and agent a^*. Further, all dummy papers $\tilde{p}_i^c$ for $i \in [n + c - 2]$ are authored by the special agent a_*^c. For a vertex $v \in V^c$, paper p_v^c is authored by the special agent a_*^c, all agents corresponding to vertices from $V^c \setminus \{v\}$ or to vertices adjacent to v in G, i.e., p_v^c is authored by agents $\{a_*^c\} \cup \{a_{v'}^c \mid v \neq v' \in V^c\} \cup \{a_{v'}^{c'} \mid c' \in [k], v' \in V^{c'}, \{v, v'\} \in E\}$. Each agent is qualified to review all papers and we set $c_{\text{reviewer}} = d_{\text{paper}} = 1$ and $z = 2$.

Lemma 1. *If the given instance $\mathcal{I}$ of* MULTICOLORED INDEPENDENT SET *is a YES-instance, then the constructed instance $\mathcal{I}'$ of* CYCLE-FREE REVIEWING *is a YES-instance.*

Proof. Let $V' = \{w^1, \ldots w^k\} \subseteq V$ be a independent set of size k in the given MULTICOLORED INDEPENDENT SET instance $\mathcal{I}$ with $w^c \in V^c$ for $c \in [k]$. From this we construct a solution for the constructed CYCLE-FREE REVIEWING instance $\mathcal{I}'$ as follows. Agent a^* reviews paper p^*. For $c \in [k]$, special agent a_*^c reviews special paper p_*^c. Vertex agents $\{a_{v'}^c \mid v' \in V^c \setminus \{w^c\}\}$ are assigned arbitrarily to dummy papers $\tilde{p}_1^c, \ldots, \tilde{p}_{n+c-2}^c$. Lastly, vertex agent $a_{w^c}^c$ reviews paper $p_{w^c}^c$ and the dummy agents from class c are assigned arbitrarily to the remaining vertex papers from this class. Note that by construction, the described assignment is 1-1 valid. Moreover, it is easy to verify that no agent reviews a paper authored by it so it remains to check for reviewing cycles of length two. All special agents are only authors of papers from their color class but review papers authored solely by agents outside their color class. Thus there exist no review cycles involving special agents. All papers a^* wrote are reviewed by special agents so a^* cannot be part of a review cycle. Dummy agents only write papers that are reviewed by special agents and a^* so no dummy agent can be part of a review cycle. Thus, every possible review cycle of length two needs to involve two vertex agents. As no dummy paper is written by a vertex agent, the only vertex agents that review papers authored by other vertex agents are those assigned to vertex papers, i.e., agents $\{a^1_{w^1}, \dots a^k_{w^k}\}$. Assume for the sake of contradiction that $a^i_{w_i}$ (which reviews paper $p^i_{w_i}$) forms a cycle with reviewer $a^{i'}_{w^{i'}}$ with $i \neq i' \in [k]$. However, from this it follows by the definition of a review cycle that $a^{i'}_{w^{i'}}$ is an author of paper $p^i_{w_i}$, which implies that $\{w^i, w^{i'}\} \in E$ contradicting that V' is an independent set. □

We now turn to proving the backwards direction of the reduction. To do this, we first identify several assignments that need to be made in all solutions to the constructed CYCLE-FREE REVIEWING instance. We start by proving that a^* needs to review p^*.

Lemma 2. *In every* 1-1 *valid* 2*-cycle-free assignment in the constructed instance* $\mathcal{I}'$*,* a^* *reviews* p^*.

Proof. Recall that all agents except all special agents and agent a^* are authors of p^*. So for the sake of contradiction let us assume that special agent a^c_* for some $c \in [k]$ reviews p^*. However, to prevent a reviewing cycle, this implies that only the remaining $k-1$ special agents and a^* can review papers written by a^c_*. However, as a^c_* is an author of all vertex papers corresponding to vertices from V^c and we have assumed that each set V^c consists of more than k vertices, these k agents are not enough to review all papers written by a^c_*, a contradiction. □

We next prove that a^c_* reviews p^c_* for all $c \in [k]$. For this, we need the following lemma:

Lemma 3. *In every* 1-1 *valid* 2*-cycle-free assignment in the constructed instance* $\mathcal{I}'$*, if* a^c_* *reviews paper* $p^{c'}_*$ *for* $c, c' \in [k]$*, then only vertex and dummy agents from class* c' *and special agents can review dummy and vertex papers from class* c*.*

Proof. Note that the special agent a^c_* is an author of all dummy and vertex papers from color class c. Moreover, paper $p^{c'}_*$ is authored by all dummy and vertex agents from color classes different from c'. Thus, if a^c_* reviews $p^{c'}_*$, then no vertex or dummy agent from a class different from c' can review papers written by a^c_*. As a^c_* authors all dummy and vertex papers from class c, the lemma follows. □

Using this, we are able to prove that each special agent reviews the corresponding special paper.

Lemma 4. *In every* 1-1 *valid* 2*-cycle-free assignment in the constructed instance* $\mathcal{I}'$*, for* $c \in [k]$*,* a^c_* *reviews* p^c_**.*

Proof. By Lemma 2, a^* is assigned to p^*, which is authored by all dummy agents and vertex agents. Thus, to prevent the existence of reviewing cycles, only special agents can review papers written by a^*. As for each $c \in [k]$, p^c_* is written by a^*, it follows that the set of k agents $\{a^c_* \mid c \in [k]\}$ needs to review the set of k papers $\{p^c_* \mid c \in [k]\}$. For the sake of contradiction, let us assume that special agent a^c_* reviews paper $p^{c'}_*$ for $c \neq c' \in [k]$. We assume without loss of generality that $c' < c$ (if there exists a pair where $a^{\tilde{c}}_*$ reviews paper $p^{\tilde{c}'}_*$ with $\tilde{c} < \tilde{c}'$ there also has to exist one with $c' < c$). By Lemma 3 and as special agents need to review special papers, from this it follows that only dummy and vertex agents from color c' can review the vertex and dummy agents from class c (which are all written by a^c_*). As we have assumed that $c' < c$, the number of these agents $(2n + 2c' - 3)$ does not suffices to review all of these papers $(2n + 2c - 3)$, a contradiction. □

We are now ready to prove the correctness of the backwards direction of the reduction:

Lemma 5. *If the constructed instance* $\mathcal{I}'$ *of* CYCLE-FREE REVIEWING *is a YES-instance, then the given instance* $\mathcal{I}$ *of* MULTICOLORED INDEPENDENT SET *is a YES-instance.*

Proof. From Lemma 2, Lemma 3, and Lemma 4 it follows that for each color $c \in [k]$ every vertex and dummy agent from this color class needs to review a vertex or dummy paper from this color class and that each vertex or dummy paper from this color class needs to be reviewed by a vertex or dummy agent from this color class. As there exist $n + c - 2$ dummy agents from color class c but $n + c - 1$ vertex papers at least one vertex paper from color class c needs to be reviewed by a vertex agent from color class c. Note that for each $v \in V^c$, agent a_v^c is an author of all vertex papers except p_v^c. Thus, for each color $c \in [k]$ there needs to exist (at least) one agent $a_{w_c}^c$ for some $w_c \in V^c$ that reviews $p_{w_c}^c$. So let $a_{w_1}^1, \cdots, a_{w_k}^k$ be a list of those agents (containing one vertex agent from each color class). We claim that $\{w_1, \ldots, w_k\}$ forms an independent set in G. For the sake of contradiction assume that $\{w_c, w_{c'}\} \in E$ for $c \neq c' \in [k]$, then by construction it follows that $a_{w_c}^c$ who reviews paper $p_{w_c}^c$ is an author of paper $p_{w_{c'}}^{c'}$ and similarly $a_{w_{c'}}^{c'}$ who reviews $p_{w_{c'}}^{c'}$ is an author of $p_{w_c}^c$. Thus, $a_{w_c}^c$ and $a_{w_{c'}}^{c'}$ form a reviewing cycle, a contradiction. □

From Lemma 1 and Lemma 5, Theorem 2 directly follows:

Theorem 2. CYCLE-FREE REVIEWING *is NP-hard even if each agent is qualified to review all papers,* $n_A = n_P$, $c_{\textit{reviewer}} = d_{\textit{paper}} = 1$, *and* $z = 2$.

The reduction from Theorem 2 heavily relies on the possibility that an agent reviews a paper written by an agent with whom she has a joint paper. As some conferences might declare an automatic conflict of interest for co-authors, we now consider the case where an agent is qualified to review all papers that are not authored by one of her co-authors:

Theorem 3. CYCLE-FREE REVIEWING *is NP-hard even if each agent is qualified to review all papers that are not written by one of her co-authors,* $c_{\textit{reviewer}} = d_{\textit{paper}} = 1$, *and* $z = 2$.

Proof. We reduce from CYCLE-FREE REVIEWING with $c_{\text{reviewer}} = d_{\text{paper}} = 1$, and $z = 2$ where agents are not qualified to review papers of co-authors, which is NP-hard as proven in Theorem 1. We assume without loss of generality that for each paper there is one agent who is not qualified to review it.

Construction. Given an instance $\mathcal{I} = ((A \cup P, E_A \cup E_P), c_{\text{reviewer}} = 1, d_{\text{paper}} = 1, z = 2)$ of CYCLE-FREE REVIEWING, we construct a new instance $\mathcal{I}'$ with agents A' and papers P' and $c'_{\text{reviewer}} = d'_{\text{paper}} = 1$ and $z' = 2$. We start by setting $A' = A$. Next, we add agents w, x, y, and z to A'. For each agent $a \in A$ and each paper $p \in P$, we insert an agent a_p to A' and add a so called *agent* paper which is authored by a and a_p to P'. For each paper $p \in P$, we introduce an agent b_p to A'. Moreover, we introduce $n_A \cdot n_P$ dummy agents $d_1, \ldots, d_{n_A \cdot n_P}$ to A'. We introduce five different (types of) papers in P':

- For each paper $p \in P$, we introduce a paper p to P' that is written by all authors of p, agent b_p and by agents a_p for all agents $a \in A$ that are not qualified to review p in $\mathcal{I}$.
- We introduce a paper q authored by w and a_p for all $a \in A$ and $p \in P$.
- We introduce a paper q' authored by x and a_p for all $a \in A$ and $p \in P$ to P'.
- We introduce a paper r authored by y and all dummy agents $d_1, \cdots, d_{n_A \cdot n_P}$ and b_p for each $p \in P$.
- We introduce a paper r' authored by z and all dummy agents $d_1, \cdots, d_{n_A \cdot n_P}$ and b_p for each $p \in P$.

Each agent is qualified to review all papers that are not written by one of her co-authors.

(⇒) Given a 1-1-valid 2-cycle-free review assignment for $\mathcal{I}$, we construct a 1-1-valid 2-cycle-free review assignment for $\mathcal{I}'$ as follows. All agents from A still review the same papers as in the given assignment (which are all still qualified to do so because we have not added or removed any papers with two authors from A apart from copies of papers from P). Agent w reviews r, agent x reviews r', y reviews q', and z reviews q (which are all are qualified to do so). Moreover, the dummy agents are assigned arbitrarily to the agent papers, which they are qualified to review because dummy agents only author papers together with agents $\{b_p \mid p \in P\}$ and y and z.

Concerning review cycles, note that agents $\{a_p, b_p \mid a \in A, p \in P\}$ do not review any paper. Moreover dummy agents only review papers written by agents $\{a_p, a \mid a \in A, p \in P\}$ but are only reviewed by x and w and thus cannot be part of a review cycle. Moreover, also no agent from A can be part of a review cycle because there was no such cycle in the given review assignment and all agents that author a paper reviewed by an agent from A are not part of a review cycle. Thus, any review cycle needs to consists of w, x, y, and z. Note that w reviews a paper of y, y reviews a paper of x, x reviews a paper of z, and z reviews a paper of w. Thus, these agents form a 4-cycle but no 2-cycle.

($\Leftarrow$) Given a 1-1-valid 2-cycle-free review assignment for $\mathcal{I}'$, we claim that this assignment restricted to agents from A and papers from P is a solution to the given instance $\mathcal{I}$. To prove this, we will argue for all agents from $A' \setminus A$ that they cannot review a paper from P from which the correctness directly follows, as we have not added any authors from A to papers from P. Fix some paper $p' \in P$. We now iterate over all agents $A' \setminus A$ and argue why they cannot review p'. As we have assumed in $\mathcal{I}$ that for all papers there is an agent not qualified to review it, it follows that p' has an author ${a^*}_{p'}$ for some $a^* \in A$ and author $b_{p'}$ in $\mathcal{I}'$. For agent w and x it holds that both have a joint paper with ${a^*}_{p'}$ and thus cannot review p'. Next, note that as, for each $a \in A$ and $p \in P$, a_p is either identical to ${a^*}_{p'}$ or has a joint paper with ${a^*}_{p'}$ none of these agents can review p'. Lastly, d_i for some $i \in [n_A \cdot n_P]$, b_p for $p \in P \setminus \{p'\}$, and y and z have a joint paper with $b_{p'}$, which is an author of p'. Thus, all these agents (and thereby no agent from $A' \setminus A$) can review p'.

□

3.3 Single-Author-Single-Paper Setting

In their theoretical analysis, Barrot et al. [2020] focus on CYCLE-FREE REVIEWING where each agent writes a single-author paper (we speak of an agent and its paper interchangeably) and qualifications are symmetric, i.e., if an agent a is qualified to review agent b, then b is qualified to review a. They prove that this problem is NP-hard for $c_{\text{reviewer}} = d_{\text{paper}} = 1$ and $z = 5$ (without bounds on Δ_A^+ or Δ_P^-) but polynomial-time solvable for arbitrary $c_{\text{reviewer}} = d_{\text{paper}}$ for $z = 2$. We close the gap between these two results and extend their general picture by proving that for $c_{\text{reviewer}} = d_{\text{paper}} = 2$, CYCLE-FREE REVIEWING is NP-hard for $z = 3$ even if qualifications are symmetric and each agent is only qualified to review four agents, i.e., we need to decide for each agent a which two of these four agents review a and which two of these agents will get a review from a.

Theorem 4. CYCLE-FREE REVIEWING *is NP-hard, even if $z = 3$, $c_{reviewer} = d_{paper} = 2$, $\Delta_A^- = \Delta_P^+ = 1$, $n_A = n_P$, each agent is qualified to review exactly four papers and if an agent a can review the paper written by agent b, then b can review the paper of a.*

To prove Theorem 4, we reduce from TWO-IN-FOUR-SATISFIABILITY, a variant of SATISFIABILITY, where given a propositional formula φ over variables X where each clause contains four different literals, the question is whether there exists an assignment α of variables X such that in each clause exactly two out of four literals are satisfied. As to the best of our knowledge, this variant of SATISFIABILITY has not been considered before, we start by proving that it is NP-hard even if each literal appears exactly twice positive and twice negative:

Proposition 2. TWO-IN-FOUR-SATISFIABILITY *is NP-hard, even if each variable appears exactly twice positive and twice negative.*

Proof. In MONOTONE NOT-ALL-EQUAL 3-SAT, we are given a propositional formula where each clause contains three different positive literals and the question is whether there is a variable assignment such that in each clause at least one literal is set to true and at least one is set to false. Reducing MONOTONE NOT-ALL-EQUAL 3-SAT to TWO-IN-FOUR-SATISFIABILITY (without any additional restrictions) is straightforward: Given an instance of MONOTONE NOT-ALL-EQUAL 3-SAT, for each clause, we introduce a new variable which we add to the clause. Thereby, we can extend a valid assignment α for the MONOTONE NOT-ALL-EQUAL 3-SAT instance by setting for a clause the newly introduced variable to true if α originally sets only one literal from this clause to true and to false if α originally sets only one literal from this clause to false. The reverse direction is immediate. However, to achieve that each variable appears twice positive and twice negative, a slightly more involved approach is needed.

In fact, for simplicity, we reduce from the NP-hard variant of MONOTONE NOT-ALL-EQUAL 3-SAT where each variable appears in exactly four clauses [Darmann and Döcker, 2020]. Given an instance $\varphi = C_1 \wedge \cdots \wedge C_m$ over variables X of MONOTONE NOT-ALL-EQUAL 3-SAT, note that m needs to be even, as there are $m = \frac{4 \cdot |X|}{3}$ and m needs to be an integer. We now construct a new propositional formula ϕ over variable set X' as follows. For each clause $C_i = w \vee x \vee y$ for $i \in [m]$, we add variables w_i, x_i, y_i, and z_i to X' and clauses $w_i \vee x_i \vee y_i \vee z_i$ and $\overline{w_i} \vee \overline{x_i} \vee \overline{y_i} \vee \overline{z_i}$ to ϕ. Now, every variable appears once negative and once positive. It remains to link the copies of each variable.

We do this for each variable separately. Let $x \in X$ be some original variable and let $j_1, \ldots, j_4$ denote the list of all clauses where x appears in φ. We introduce dummy variables a_x^1 and a_x^2 to X' and add clauses $\overline{x_{j_1}} \vee x_{j_2} \vee a_x^1 \vee \overline{a_x^1}$, $\overline{x_{j_2}} \vee x_{j_3} \vee a_x^1 \vee \overline{a_x^1}$, $\overline{x_{j_3}} \vee x_{j_4} \vee a_x^2 \vee \overline{a_x^2}$, and $\overline{x_{j_4}} \vee x_{j_1} \vee a_x^2 \vee \overline{a_x^2}$ to ϕ. As for each $j \in [3]$, exactly one of $\overline{x_{j_i}}$ and $x_{j_{i+1}}$ need to be set to true, these clauses enforce that x_{j_1}, x_{j_2}, x_{j_3}, and x_{j_4}, all have the same truth value. Lastly, for $i \in [\frac{m}{2}]$, we add twice the clause $z_{2i-1} \vee \overline{z_{2i-1}} \vee z_{2i} \vee \overline{z_{2i}}$ which are always trivially satisfied.

The correctness of the reduction is immediate and all variables appear twice positive and twice negative in ϕ. $\square$

Using this, we are now ready to prove Theorem 4:

Proof of Theorem 4. We reduce from TWO-IN-FOUR-SATISFIABILITY where each variable appears exactly twice positive and twice negative.

Construction. Given an instance of TWO-IN-FOUR-SATISFIABILITY consisting of a propositional formula $\varphi = C_1 \wedge \cdots \wedge C_m$ over variables $X = \{x_1, \ldots, x_n\}$, for $i \in [n]$, we denote as $t_{i,1}^{\text{pos}}$ and $t_{i,2}^{\text{pos}}$ the indices of the two clauses in which variable x_i appears positive and as $t_{i,1}^{\text{neg}}$ and $t_{i,2}^{\text{neg}}$ the indices of the two clauses in which variable x_i appears negative. From this, we construct an instance of PEER CYCLE-FREE REVIEWING as follows. For $i \in [n]$, we introduce four agents a_i^{pos}, a_i^{neg}, a_i^1, and a_i^2 (constituting a gadget modeling this variable). Moreover, for $j \in [m]$, we introduce one agent b_j. Qualification are symmetric, i.e., if agent a is qualified to review b, then b is qualified to review a. For $i \in [n]$, a_i^{pos} is qualified to review a_i^1, a_i^2, $b_{t_{i,1}^{\text{pos}}}$, and $b_{t_{i,2}^{\text{pos}}}$ (and the other way round). Moreover, a_i^{neg} is qualified to review a_i^1, a_i^2, $b_{t_{i,1}^{\text{neg}}}$, and $b_{t_{i,2}^{\text{neg}}}$ (and the other way round). Lastly, a_i^1 and a_i^2 are qualified to review each other and a_i^2 and a_{i+1}^1 are qualified to review each other (where i is taken modulo n). We set $c_{\text{reviewer}} = d_{\text{paper}} = 2$ and $z = 3$.

($\Rightarrow$) Let α be an assignment of variables in X that is a solution to the given TWO-IN-FOUR-SATISFIABILITY instance. For $i \in [n-1]$, we let a_i^1 review a_i^2 and a_i^2 review a_{i+1}^1. Moreover, we let a_n^1 review a_n^2 and a_n^2 review a_1^1.

For $i \in [n]$ where x_i is set to true by α, we let a_i^1 and a_i^2 review a_i^{pos} and a_i^{neg} review a_i^1 and a_i^2. Moreover we let a_i^{pos} review $b_{t_{i,1}^{\text{pos}}}$ and $b_{t_{i,2}^{\text{pos}}}$ and we let $b_{t_{i,1}^{\text{neg}}}$ and $b_{t_{i,2}^{\text{pos}}}$ review a_i^{neg}. Conversely, for $i \in [n]$

where x_i is set to false by α, we let a_i^1 and a_i^2 review a_i^{neg}, we let a_i^{pos} review a_i^1 and a_i^2. Moreover, we let a_i^{neg} review $b_{t^{\text{neg}}_{i,1}}$ and $b_{t^{\text{neg}}_{i,2}}$ and let $b_{t^{\text{pos}}_{i,1}}$ and $b_{t^{\text{pos}}_{i,2}}$ review a_i^{pos}.

As α sets exactly two literals in each clause to true and two to false, for each $j \in [m]$, b_j is reviewed by two agents and reviews two agents. The same also holds for all other agents, implying that the constructed review assignment is 2-2 valid. It is easy to see that there are no 2-cycles. Moreover, as no two agents b_i and b_j for $i \neq j$ are qualified to review each other and, for no $j \in [m]$, are there two agents that are both qualified to review b_j and that are qualified to review each other, each 3-cycle needs to solely consist of agents from a gadget corresponding to a single variable. So let us fix some $i \in [n]$. The only possible 3-cycles consist of a_i^{pos}, a_i^1 and a_i^2 or a_i^{neg}, a_i^1 and a_i^2. However, there is no such 3-cycle, as either a_i^{pos} reviews both a_i^1 and a_i^2 and both a_i^1 and a_i^2 review a_i^{neg}, or a_i^{neg} reviews both a_i^1 and a_i^2 and both a_i^1 and a_i^2 review a_i^{pos}. Thus, the constructed assignment is 3-cycle-free.

($\Leftarrow$) Assume we are given a 2-2-valid 3-cycle-free review assignment in the constructed CYCLE-FREE REVIEWING instance. Assume that a_n^2 reviews a_1^1 in the given assignment (if a_1^1 reviews a_n^2 an analgous argument works). We now argue that a_1^1 needs to review a_1^2. Assume for the sake of contradiction that this is not the case, then as a_1^1 is reviewed by a_n^2 and a_1^2, she needs to review a_1^{pos} and a_1^{neg}. However, to prevent a 3-cycle, a_1^2 then needs to review a_1^{pos} and a_1^{neg}, a contradiction (as a_1^2 gives three reviews). Next, we want to argue that a_1^2 reviews a_2^1. For the sake of contradiction, assume that a_2^1 reviews a_1^2. Then, a_1^2 already gets two reviews and thus needs to review a_1^{neg} and a_1^{pos}. However, as a_1^1 already reviews a_1^2 either a_1^{neg} or a_n^{pos} review a_1^1 which leads to a 3-cycle together with a_1^1 and a_1^2. Applying the same arguments inductively, it follows that for $i \in [n-1]$, a_i^1 review a_i^2 and a_i^2 reviews a_{i+1}^1 and that a_n^1 reviews a_n^2.

Further, observe that for each $i \in [n]$ agents a_i^1 and a_i^2 either both review a_i^{pos} or both get reviews from a_i^{pos}. For the sake of contradiction, assume that this is not the case. If a_i^2 reviews a_i^{pos} and a_i^{pos} reviews a_i^1, then we have a 3-cycle consisting of these three agents. Otherwise, a_i^1 reviews a_i^{pos} and a_i^{pos} reviews a_i^2. However, as the given assignment is 2-2-valid, from this it follows that a_i^2 reviews a_i^{neg} and a_i^{neg} reviews a_i^1, which leads to a 3-cycle consisting of a_i^1, a_i^2, and a_i^{neg}. Thus, we have reached a contradiction proving our initial claim. Moreover, as the given assignment is 2-2 valid, in case that a_i^1 and a_i^2 both review a_i^{pos}, then a_i^{neg} reviews a_i^1 and a_i^2, and in case that a_i^{pos} reviews both a_i^1 and a_i^2, then a_i^1 and a_i^2 both review a_i^{neg}. We now construct an assignment α by, for $i \in [n]$, setting variable x_i to true if a_i^1 and a_i^2 review a_i^{pos} and x_i to false if a_i^1 and a_i^2 review a_i^{neg}. Using our argument from above, it follows that α is well-defined. Moreover, the given assignment is 2-2-valid, if α sets a literal to true, then the agents corresponding to this literal review the agents corresponding to the two clauses in which the literal appears. Similarly, if α sets a literal to false, then the agents corresponding to this literal get a review from the two agents corresponding to the two clauses in which the literal appears. Thus, as each agent corresponding to a clause gets and issues two reviews (as the given assignment is 2-2-valid), it follows that α sets for each clause exactly two literals to true and thus that α is a solution to the given instance of TWO-IN-FOUR-SATISFIABILITY. □

4 Polynomial-Time Solvable Special Cases

In this section, we identify conditions under which a short-cycle-free review assignment provably exists and can be computed in polynomial time. As we will see in our experiments, the subsequently presented algorithms provide short-cycle-free review assignments even beyond the theoretical limitations we discuss below. As we are interested in computing z-cycle-free review assignments for $z \geq 1$, no author is allowed to review one of its own papers. That is why throughout this section we assume that we do not have $(a,p) \in E_A$ and $(p,a) \in E_P$ at the same time.

Our algorithms in this section are based on the following simple observation: Given a partial z-cycle-free review assignment E' and a paper $p \in P$ that requires more assigned reviewers, the number

Algorithm 1: A greedy algorithm computing a d_{paper}-d_{paper}-valid completely cycle-free assignment E'.

1 $E' \leftarrow \emptyset$; $S_0 \leftarrow$agents without papers
  /* $\phi_i(a)$ is the free reviewing capacity of a before iteration i of the for-loop from Line 3; each agent reviews at most d_{paper} papers */
2 **foreach** $a \in A$ **do** $\phi_0(a) \leftarrow d_{\text{paper}}$
  /* Assign reviewers to one paper per iteration: */
3 **for** $i \leftarrow 0$ **to** $n_P - 1$ **do**
4   **foreach** $a \in A$ **do** $\phi_{i+1}(a) \leftarrow \phi_i(a)$
5   select some $(p, a) \in E_P$ where p has no reviews yet
    /* collect qualified reviewers and assign d_{paper} of them to p */
6   $R \leftarrow \{b \in S_i \mid (b, p) \in E_A\}$
7   **for** $j \leftarrow 1$ **to** d_{paper} **do**
8   arbitrary $b \in R$ reviews p: $E' \leftarrow E' \cup \{(b, p)\}$
9     $\phi_{i+1}(b) \leftarrow \phi_i(b) - 1$; $R \leftarrow R \setminus \{b\}$
    /* collect possible reviewers for next paper */
10   $S_{i+1} := \{a\} \cup \{b \in S_i \mid \phi_{i+1}(b) > 0\}$
11 **return** E'

of potential reviewers that would create a z-cycle–if assigned to review p–is bounded by a function in z, the maximum number Δ_P^+ of authors per paper, and the maximum number c_{reviewer} of reviews per agent; the precise function is given in the subsequent proofs. Thus, assuming that the minimum number δ_P^- of potential reviewers for each paper is large compared to z, Δ_P^+, d_{paper}, and c_{reviewer}, for each paper p there are always reviewers that can be assigned to review p without creating a z-cycle. Note that in practice we can expect that z, d_{paper}, and c_{reviewer} are quite small. Moreover, while the minimum number of fitting reviewers might be not very large, it is not uncommon to assign papers to reviewers that are not "perfect". Thus, interpreting δ_P^- as the number of community members that do not have a conflict of interest actually yields relative large values for δ_P^- in practice.

We start with a very restrictive setting and then, step by step, generalize the approach and the results. First, each paper is written by exactly one author, each agent has at most one paper and we want a completely cycle-free review assignment (i. e., z-cycle-free for every $z \in \mathbb{N}$). This of course implies that some agents cannot be authors of papers and so the number n_P of papers is smaller than the number n_A of agents. However, it allows Algorithm 1 to work (implicitly) with the topological ordering of the (acyclic) review assignment while constructing it.

Proposition 3. *If $\Delta_A^- \leq 1 = \delta_P^+ = \Delta_P^+$, $d_{paper} \leq c_{reviewer}$, and $\delta_P^- \geq n_P + d_{paper}$, then Algorithm 1 computes a d_{paper}-d_{paper}-valid and completely cycle-free review assignment in linear time.*

Proof. We first show the correctness of Algorithm 1. Clearly, if in each iteration of the loop in Line 3 the set of eligible reviewers R (see Line 6) is of size at least d_{paper}, then a completely cycle-free review assignment is created as each agent only reviews papers from agents "occurring" later during the algorithm. Observe that if $|S_i| \geq n_A - \delta_P^- + d_{\text{paper}}$ for $i \in \{0, \ldots, n_P - 1\}$, then in iteration i we have $|R| \geq d_{\text{paper}}$: There are at most $n_A - \delta_P^-$ agents in S_i that cannot review p (the corresponding edge is not in E_A) and, thus, at least d_{paper} agents in S_i are eligible to review p. It remains to show that $|S_i| \geq n_A - \delta_P^- + d_{\text{paper}}$ for all $i \in \{0, \ldots, n_P - 1\}$ follows from our assumptions. By assumption of the lemma we have $n_P \leq \delta_P^- - d_{\text{paper}}$. Hence, $|S_0| = n_A - n_P \geq n_A - \delta_P^- + d_{\text{paper}}$. We next show that $|S_i| \geq |S_0|$ for all $i \in [n_P - 1]$. Observe that at the start we have $\sum_{a \in S_0} \phi_0(a) = |S_0| \cdot d_{\text{paper}}$. Moreover, after the ith iteration of the loop in Line 3 we have $\sum_{a \in S_{i+1}} \phi_{i+1}(a) = \sum_{a \in S_i} \phi_i(a)$ as each paper gets d_{paper} reviews

Algorithm 2: Greedy algorithm to compute a c_{reviewer}-d_{paper}-valid z-cycle-free review assignment E'.

1 $E' \leftarrow \emptyset$
2 **while** $\exists p \in P \colon |N^-(p, E')| < d_{paper}$ **do**
3   **if** $\exists (a,p) \in E_A \setminus E' \colon E' \cup \{(a,p)\}$ *is z-cycle free and* $|N^+(a, E')| < c_{reviewer}$ **then**
        `/* Case 1: greedy assignment of reviews as long as no z-cycles are created: */`
4         $E' \leftarrow E' \cup \{(a,p)\}$
5   **else**
        `/* Case 2: replace one review assignment by two: */`
6   pick $(a', p') \in E'$ and $a \in A$ so that $|N^+(a, E')| < c_{\text{reviewer}}$, $(a', p), (a, p') \in E_A$, and $(E' \setminus \{(a', p')\}) \cup \{(a', p), (a, p')\}$ is z-cycle free
7         $E' \leftarrow (E' \setminus \{(a', p')\}) \cup \{(a', p), (a, p')\}$
8 **return** E'

and the reviewer a in $S_{i+1} \setminus S_i$ starts with $\phi_{i+1}(a) = d_{\text{paper}}$. Observe that $\phi_i(a) \leq d_{\text{paper}}$ for all $a \in A$ and $i \in \{0, \ldots, n_P - 1\}$. Thus, we have $|S_i| d_{\text{paper}} \geq \sum_{a \in S_i} \phi_i(a) = \sum_{a \in S_0} \phi_0(a) = |S_0| d_{\text{paper}}$ and, hence, $|S_i| \geq |S_0|$. This completes the correctness proof.

As to the running time, everything outside the loop starting in Line 3 clearly runs in linear time. As to the part inside the loop, note that by keeping just one array of length n_A we can store the values of ϕ in linear time. Moreover, the reviewers for p are selected arbitrarily from R, which is doable in $|N^-(p, E_A)|$ time. Hence, the loop in Line 3 can be processed in $O(n_P + |E_A|)$ time. Thus, the overall algorithm runs in $O(n_A + n_P + |E_A|)$, that is, linear time. $\square$

For our next result we replace the completely cycle-free property of the resulting review assignment with z-cycle freeness. This implies that the idea of constructing the review assignment along its topological ordering (as done by Algorithm 1) cannot be employed. Instead, Algorithm 2 constructs greedily a maximal z-cycle-free assignment and then extends the assignment by replacing one review assignment by two other assignments. The argument behind the replacement strategy is an extension of the argument in Algorithm 1 that there are always enough reviewers to assign in Lines 7 to 9.

To keep our arguments simple we first consider the case that each agent reviews at most one paper and each paper requires one review. Moreover, as before, we are in the setting that each paper has one author and each agent authors at most one paper. Formally, we have the following.

Proposition 4. *If* $\Delta_A^- \leq 1 = \delta_P^+ = \Delta_P^+ = c_{reviewer} = d_{paper}$, $n_A \geq n_P$, $\delta_A^+ > z$, $\delta_P^- > z$, *and* $n_P \leq \delta_A^+ + \delta_P^- - 2z$, *then Algorithm 2 computes a* $c_{reviewer}$-d_{paper}*-valid z-cycle-free review assignment in polynomial time.*

Proof. Obviously, Algorithm 2 terminates after at most n_P iterations of the while loop as in each iteration the number of assigned reviews increases. Moreover, a c_{reviewer}-d_{paper}-valid z-cycle-free review assignment is returned if a, a', p' as described in case 2 (Line 6) always exist. To prove their existence, we introduce some notation. For some $v \in A \uplus P$ let $N_z^+(v, E' \uplus E_P)$ be the z-out-neighborhood of v, that is, the set of vertices that can be reached from v in the review graph $(A \uplus P, E' \uplus E_P)$ via a path of length at least one and at most $2z$. Similarly, let $N_z^-(v, E' \uplus E_P)$ be the $2z$-in-neighborhood of v, that is, the set of vertices that can reach v in the review graph $(A \uplus P, E' \uplus E_P)$ via a path of length at least one and at most $2z$. Note that if $v \in N_z^-(v, E' \uplus E_P)$, then also $v \in N_z^+(v, E' \uplus E_P)$ and v is contained in a review cycle of length z (that is a directed cycle of length $2z$ in $(A \uplus P, E' \uplus E_P)$). Subsequently, we present upper bounds on the size of $N_z^-(v, E' \uplus E_P)$ and $N_z^+(v, E' \uplus E_P)$ for $v \in A \uplus P$ thereby proving the existence of a, a', p'.

Let $p \in P$ be the paper without reviewer selected in Line 2 when the algorithm enters case 2. Let $A_p \subseteq A$ be the set of agents that could review p without creating a z-cycle, that is, $A_p := \{a \in A \mid (a,p) \in E_A \wedge E' \cup \{(a,p)\}$ is z-cycle free$\}$. Since $d_{\text{paper}} = c_{\text{reviewer}} = \Delta_P^+ = 1$, there are at most z agents whose assignment to review p would create a review cycle, that is, $|N_z^+(p, E' \uplus E_P) \cap A| \leq z$, and thus $|A_p| \geq \delta_P^- - z$. Since we are in case 2, no more review assignments could be added without creating a z-cycle. Hence, the algorithm assigned the at least $\delta_P^- - z$ potential reviewers in A_p to different papers. Let P_p be the set of these papers. Since $d_{\text{paper}} = c_{\text{reviewer}} = 1$ we have $|P_p| = |A_p| \geq \delta_P^- - z$.

Let $a \in A$ be an arbitrary agent without assigned review, that is, $\nexists p''\colon (a, p'') \in E'$. Since $d_{\text{paper}} = c_{\text{reviewer}} = \Delta_A^- = 1$, we have $|N_z^-(a, E' \uplus E_P) \cap P| \leq z$. Thus, there are $\delta_A^+ - z$ papers that a could review without creating a z-cycle; let P_a denote the set of these papers. Since we assume that $n_P \leq \delta_A^+ + \delta_P^- - 2z$, it follows that there is a $p' \in P_a \cap P_p$. By definition of P_p there is an agent a' with $(a', p') \in E'$ and $a' \in A_p$. Thus, a, a', p' exist and E' can be updated to $(E' \setminus \{(a', p')\}) \cup \{(a', p), (a, p')\}$ in Line 7. $\square$

We now turn our attention to our general case where agents can author and review many papers and papers can have multiple authors and can require several reviews. While the conditions that guarantee the existence of a z-cycle-free review assignment need adjustments, we can still use Algorithm 2 together with a correctness proof that follows a similar pattern as the proof of Proposition 4.

Theorem 5. *If, $n_A \cdot c_{\text{reviewer}} \geq n_P \cdot d_{\text{paper}}$, $\delta_A^+ > 2(\Delta_A^- \cdot d_{\text{paper}})^z + c_{\text{reviewer}}$, $\delta_P^- > 2(\Delta_P^+ \cdot c_{\text{reviewer}})^z + d_{\text{paper}}$, and $n_P \leq \delta_A^+ - 2(\Delta_A^- \cdot d_{\text{paper}})^z - c_{\text{reviewer}} + (c_{\text{reviewer}}/d_{\text{paper}})(\delta_P^- - 2(\Delta_P^+ \cdot c_{\text{reviewer}})^z - d_{\text{paper}})$, then Algorithm 2 computes a c_{reviewer}-d_{paper}-valid z-cycle-free review assignment in polynomial time.*

Proof. We use the same notation as in the proof of Proposition 4 and similarly to this proof we need to show that a, a', p' as described in Line 6 actually always exist.

Let $p \in P$ be the paper with a missing review selected in Line 2 and the algorithm entered case 2. Let $A_p \subseteq P$ be the set of agents that could review p without creating a z-cycle, that is, $A_p := \{a \in A \mid (a,p) \in E_A \setminus E' \wedge E' \cup \{(a,p)\}$ is z-cycle free$\}$. As every paper has at most Δ_P^+ authors and every author has at most c_{reviewer} assigned papers to review, it follows that

$$\begin{aligned}|N_z^+(p, E' \uplus E_P) \cap A| &\leq \Delta_P^+ \cdot \sum_{i=0}^{z-1} (\Delta_P^+ \cdot c_{\text{reviewer}})^i \\ &= \Delta_P^+ \cdot ((\Delta_P^+ \cdot c_{\text{reviewer}})^z - 1)/((\Delta_P^+ \cdot c_{\text{reviewer}}) - 1) \\ &< 2(\Delta_P^+ \cdot c_{\text{reviewer}})^z.\end{aligned}$$

Thus, $|A_p| > \delta_P^- - 2(\Delta_P^+ \cdot c_{\text{reviewer}})^z - d_{\text{paper}}$, as at most $d_{\text{paper}} - 1$ agents are already assigned to p and at most $|N_z^+(p, E' \uplus E_P) \cap A| < 2(\Delta_P^+ \cdot c_{\text{reviewer}})^z$ agents cannot review p because this would cause a review cycle of length at most z. When case 2 was entered, no more review assignment could be added without creating a z-cycle. Hence, the algorithm assigned the potential reviewers in A_p already to different papers. Let P_p be the set of these papers. Note that $|P_p| \geq c_{\text{reviewer}}|A_p|/d_{\text{paper}}$.

Let $a \in A$ be an arbitrary agent that can do one more review, that is, $|N^+(a, E')| < c_{\text{reviewer}}$. Using a similar argument as above, we can show $|N_z^-(a, E' \uplus E_P) \cap P| < 2(\Delta_A^- \cdot d_{\text{paper}})^z$. Thus, there are more than $\delta_A^+ - 2(\Delta_A^- \cdot d_{\text{paper}})^z - c_{\text{reviewer}}$ papers that a could review additionally without creating a z-cycle; let P_a denote the set of these papers. Note that by our assumptions that $\delta_A^+ > 2(\Delta_A^- \cdot d_{\text{paper}})^z + c_{\text{reviewer}}$ and $\delta_P^- > 2(\Delta_P^+ \cdot c_{\text{reviewer}})^z + d_{\text{paper}}$, P_a and P_p are both non-empty. Since $n_P \leq \delta_A^+ - 2(\Delta_A^- \cdot d_{\text{paper}})^z - c_{\text{reviewer}} + (c_{\text{reviewer}}/d_{\text{paper}})(\delta_P^- - 2(\Delta_P^+ \cdot c_{\text{reviewer}})^z - d_{\text{paper}}) < |P_a| + |P_p|$, it follows that there is a $p' \in P_a \cap P_p$. By definition of P_p there is an agent a' with $(a', p') \in E'$ and $a' \in A_p$. Thus, a, a', p' exist and E' can be updated to $(E' \setminus \{(a', p')\}) \cup \{(a', p), (a, p')\}$ in Line 7. This finishes the correctness proof. $\square$

To simplify the statement of Theorem 5 consider a "symmetric" case where $n_A \geq n_P$, $\delta_P^- = \delta_A^+$, and $\Delta_P^+ = \Delta_A^-$. For brevity, set $n := n_P$, $\delta := \delta_P^-$, and $\Delta := \Delta_P^+$. Let coi be the maximum number of papers any agent is not qualified to review/has a conflict of interest with, that is, $\mathrm{coi} = n - \delta$. Setting $c_{\text{reviewer}} = 6$ and $d_{\text{paper}} = 3$ as in our experiments we get:

Corollary 2. *If $n - 6 \geq 1.5 \cdot \mathrm{coi} + \Delta^z(6^z \cdot 2 + 3^z)$), then there always exists a 6-3-valid z-cycle-free review assignment that can be found in polynomial time.*

Considering that AAAI'22 had 9,251 submissions and that there was a submission limit of 10 papers per author and assuming that each paper has at most ten authors (implying that $\Delta = 10$) and that each author has at most 700 conflict of interests, it follows that there is a 6-3-valid 2-cycle-free review assignment computable with Algorithm 2.

As we see in the experiments in the next section, our algorithm returns $2/3/4$-cycle-free review assignments even well beyond the theoretical guarantees given above. We also remark that Algorithm 2 allows for an easy extension to the weighted case which we use in our experiments in the next section. To this end, in the first case (Line 4) we do not pick an arbitrary edge (a, p) but a eligible edge of maximum weight to be added to the assignment E'.

5 Experiments

In this section, we compare the weight of review assignments computed using different methods and analyze the occurrences of review cycles.[2] For this, we use a dataset from the 2018 International Conference on Learning Representations (ICLR '18) prepared by Xu et al. [2019]. Xu et al. [2019] collected all 911 papers submitted to ICLR '18 and the identity of all 2428 authors. As reviewers identities are unknown, they considered all authors to be reviewers and computed for each author-paper pair a similarity score.[3]

From the dataset of Xu et al. [2019], we created multiple instances of WEIGHTED CYCLE-FREE REVIEWING as follows. Given a number n_P of papers and a ratio r_{AP} of the numbers of agents and papers, we sample a subset of n_P of the 911 ICLR '18 papers and set this as our set of papers. Subsequently, we compute the set of all authors of one of these papers and sample a subset of $r_{AP} \cdot n_P$ authors and set this as our set of agents. Notably, the created instances can be seen as particularly challenging when it comes to avoiding review cycles, as in reality also "uncritical" reviewers, i.e., reviewers that do not author any paper, exist.

As done in other papers using the same dataset, we focus on the case with $d_{\text{paper}} = 3$ and $c_{\text{reviewer}} = 6$, i.e., every paper needs exactly three reviews and each agent can review at most six papers [Jecmen et al., 2020, Xu et al., 2019]. We consider three different types of review assignments: As "optimal" we denote a maximum-weight c_{reviewer}-d_{paper}-valid review assignment. Such an assignment can be computed using a simple Linear Program (LP) as, for instance, described by Taylor [2008]. As "optimal z-cycle free" we denote a maximum-weight c_{reviewer}-d_{paper}-valid z-cycle-free review assignment. This solution can be computed by treating the LP of Taylor [2008] as an Integer Linear Program (ILP) and adding for each possible i-cycle for $i \in [z]$ a separate constraint which imposes that at least one of the agent-paper pairs from the cycle is not assigned. We solved all (I)LPs using Gurobi Optimization, LLC [2021]. Lastly, as "heuristic z-cycle free", we denote a c_{reviewer}-d_{paper}-valid z-cycle-free review assignment computed by the weighted variant of Algorithm 2 as described at the end of Section 4.[4] In all experiments conducted in this section, the heuristic always returned a solution despite the fact that most of the time we are beyond the setting in which Theorem 5 guarantees this behavior of the heuristic. In experiment I presented in

[2] The code for our experiments is available at github.com/n-boehmer/Combating-Collusion-Rings-is-Hard-but-Possible.

[3] To the best of our knowledge, in all other publicly available datasets, there are similarity scores for reviewer-paper pairs but the link between the identities of authors and reviewers is missing (as this is considered sensitive information).

[4] We could not use the heuristics of Guo et al. [2018] as these are not available and their algorithm details are ambiguous.

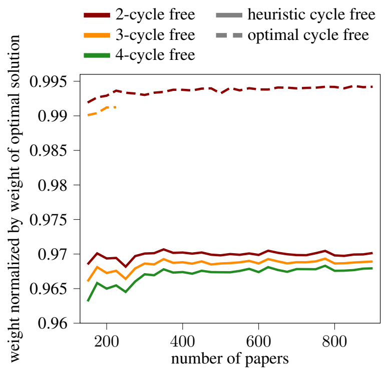

Figure 1: For different values of z, weight of an optimal/heuristic z-cycle-free assignment divided by the weight of an optimal assignment.

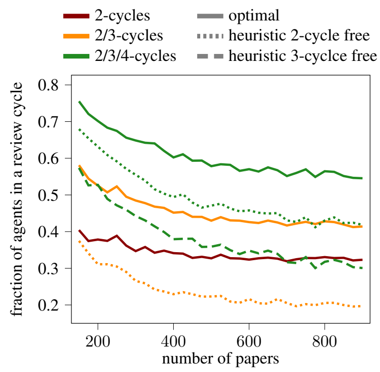

Figure 2: Fraction of agents that are part of a review cycle of at most some length for different types of assignment.

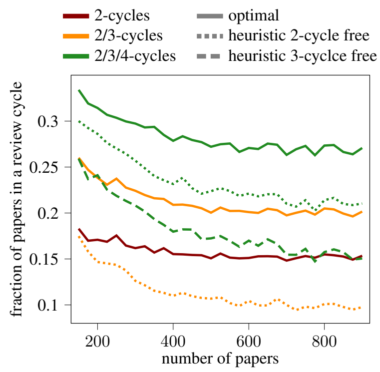

Figure 3: Fraction of papers that are part of a review cycle of at most some length for different types of assignment.

the following subsection, for $z = 2/3/4$, an unoptimized implementation of our heuristic was always able to find a z-cycle-free review assignment in less than 30 seconds, being on average around 2 times faster than the "optimal" LP, on average around 3.7 times faster than the "optimal 2-cycle free" ILP, and on average more than 100 times faster than the "optimal 3-cycle free" ILP.

5.1 Experiment I

In this experiment, we focus on the case where the total number of needed reviews is the same as the total number of reviews that can be written, which is in some sense the most "challenging" but probably also one of the more realistic scenarios. Specifically, for $n_P \in \{150, 175, \ldots, 900\}$, we prepared 100 instances with $r_{AP} = 0.5$ as described above and computed for each of these instances the optimal, heuristic 2/3/4-cycle-free, and optimal 2-cycle-free review assignment. Moreover, for all instances with $n_P \leq 225$, we also computed the optimal 3-cycle-free review assignment (for larger instances the ILP solver run out of memory.)

To measure the "price of z-cycle freeness", in Figure 1, we display the weights of different cycle-free review assignments divided by the weight of an optimal review assignment. What stands out here is that forbidding the existence of 2-cycles only comes at the cost of decreasing the assignment's weight by on average at most 0.8% (if the optimal 2-cycle-free assignment is used). Turning to the results produced by our heuristic, the quality decrease for 2/3/4-cycle-free assignments lies, on average, around 3.1%, 3.2%, and 3.3%. The weight of assignments computed using our heuristic is thus clearly worse than the weight of the optimal cycle-free assignment, yet still not far away from the the weight of an optimal assignment. What is particularly surprising here is that for both our heuristic and the optimal cycle-free assignment, whether 2, 3 or 4 cycles are forbidden seems to be rather irrelevant for the quality decrease. All in all, it is encouraging that 2/3/4-cycle freeness can be realized at a low cost independent of whether our heuristic or an ILP is used.

The necessity of dealing with review cycles is underlined by the data displayed in Figure 2. Here, we show the fraction of agents that are contained in at least one review cycle of some length in an optimal assignment and in a heuristic $2/3$-cycle-free assignment. Overall, as the number of papers increases the fraction of agents contained in review cycles constantly decreases, yet for all considered values of n_P the results are worrisome. In the optimal assignment for 150 papers, the fraction of agents contained in a review cycle of length at most $2/3/4$ is, on average, $40\%/58\%/76\%$, while even for 900 papers, still $32\%/41\%/55\%$ of agents are contained in a review cycle. Considering heuristic z-cycle-free as-

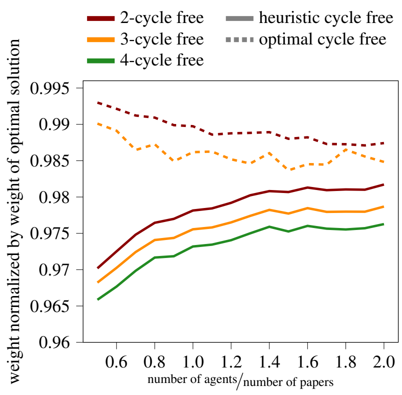

Figure 4: For different values of z, weight of a optimal/heuristic z-cycle-free assignment divided by the weight of an optimal assignment.

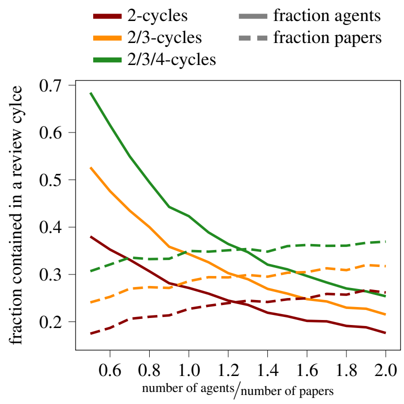

Figure 5: Fraction of agents/papers that are part of a review cycle of at most some length in an optimal assignment for 200 papers and between 100 and 400 agents.

signments, the fraction of agents contained in a cycle of length $z + 1$ is considerably lower than for the optimal solution but still non-negligible (the results for optimal $2/3$-cycle-free assignments are similar to the displayed results for our heuristic).

We also computed the fraction of papers that are contained in at least one review cycle (see Figure 3). The results are as in Figure 2 with all values roughly halved, e.g, even in the optimal assignment for 900 papers, $15\%/20\%/27\%$ of papers are contained in a review cycle of length at most $2/3/4$. An intuitive explanation for this difference between agents and papers is that the number of papers is twice the number of agents and that there exist some papers without reviewing authors. Overall, it is striking that even for a high number of papers, in an optimal assignment around 15% of papers could have a considerably higher chance of getting accepted if two agents coordinate to give each others paper better reviews and 32% of reviewers would have an opportunity to participate in such a collusion.

5.2 Experiment II

In this experiment, we analyze how the results from experiment I depend on the assumption that the supply and demand of reviews exactly matches. In particular, as describe before, for $r_{AP} \in \{0.5, 0.6, \ldots, 1.9, 2\}$ we prepared 100 instances with 200 papers and $r_{AP} \cdot 200$ agents (we also repeated this experiment for 400 and 600 papers producing similar results) and computed the different types of review assignments. Considering the assignment weights (see Figure 4), increasing r_{AP} from 0.5 to 2, the normalized weight of an optimal $2/3$-cycle-free assignment decreases by 0.005 to $0.987/0.985$, while the normalized weight of a heuristic $2/3/4$-cycle-free assignment increases by 0.01 to $0.982/0.979/0.976$: our heuristic performs particularly well if there are (considerably) more reviews available then needed; this supports our theoretical statements for our heuristic in Section 4.

Turning to the possible impact of review cycles, we visualize the fraction of agents/papers contained in a review cycle in an optimal assignment in Figure 5.[5] While the fraction of agents contained in a review cycle constantly and significantly decreases if more and more agents are added, the fraction of papers contained in a cycle constantly increases. The former observation is quite intuitive, as when more and more agents are added, the average review load decreases and even if the number of review cycles remains the same, it is likely that the fraction of agents contained in one gets smaller. The latter observation is less intuitive but probably a consequence of the fact that, starting with $r_{AP} = 0.5$, for

[5]For readability, we do not display the values for the optimal/heuristic cycle-free assignment, as their relationship to the optimal assignment is again similar as in Figure 2.

some papers none of the authors is part of the agent set, implying that these papers cannot be part of a review cycle; however, if we start to add more and more agents, more and more papers can potentially be part of a review cycle. Overall, it might be quite counter intuitive that adding more and more reviewers (that are also authors) to the reviewer pool does not decrease the number of papers contained in a review cycle but increases them.

6 Conclusion

Our work provides a first systematic analysis of CYCLE-FREE REVIEWING. On the theoretical side, we show that CYCLE-FREE REVIEWING is a computationally hard problem even in very restricted settings, yet practically relevant polynomial-time solvable special cases exist. In our practical analysis, we could show that in assignments that do not care for review cycles a high fraction of authors and papers will likely be part of a short review cycle. While collusion rings can certainly also emerge without the existence of review cycles, for example, when authors coordinate over multiple conferences [Littman, 2021, Shah, 2021], allowing so many easy opportunities means to leave a huge door unlocked without good reason: Our heuristic significantly improves the situation, since it seems to always find cycle-free review assignment at a very low quality loss.

For future work, it would be valuable to further investigate the limits of our heuristic. While our current bounds are certainly not tight, there are also clear limitations for possible extensions imposed by our NP-hardness results in quite restrictive settings from Section 3. However, a concrete and practically very relevant open question is whether the minimum degree in our analysis can be replaced by the average degree; this would make the results much more robust against outliers. Finally, due to the lack of data, we tested our model on just one dataset. Obtaining more data to test our and other models on would be extremely valuable.

Acknowledgments

NB was supported by the DFG project MaMu (NI 369/19) and by the DFG project ComSoc-MPMS (NI 369/22). RB was partially supported by the DFG project AFFA (BR 5207/1 and NI 369/15). This work was started at the research retreat of the TU Berlin Algorithms and Computational Complexity group held in September 2020.

References

Nathanaël Barrot, Sylvaine Lemeilleur, Nicolas Paget, and Abdallah Saffidine. Peer reviewing in participatory guarantee systems: Modelisation and algorithmic aspects. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS '20)*, pages 114–122. IFAAMAS, 2020.

Piotr Berman, Marek Karpinski, and Alex D. Scott. Approximation hardness of short symmetric instances of MAX-3SAT. *ECCC*, (49), 2003.

Andreas Darmann and Janosch Döcker. On a simple hard variant of not-all-equal 3-sat. *Theor. Comput. Sci.*, 815:147–152, 2020.

Naveen Garg, Telikepalli Kavitha, Amit Kumar, Kurt Mehlhorn, and Julián Mestre. Assigning papers to referees. *Algorithmica*, 58(1):119–136, 2010.

Judy Goldsmith and Robert H. Sloan. The AI conference paper assignment problem. In *Proceedings of the 22nd AAAI Conference Workshop on Preference Handling for Artificial Intelligence (MPREF '07)*, pages 53–57. AAAI Press, 2007.

Longhua Guo, Jie Wu, Wei Chang, Jun Wu, and Jianhua Li. K-loop free assignment in conference review systems. In *Proceedings of the 2018 International Conference on Computing, Networking and Communications (ICNC '18)*, pages 542–547. IEEE Computer Society, 2018.

Gurobi Optimization, LLC. Gurobi Optimizer Reference Manual, 2021. URL https://www.gurobi.com.

Pavol Hell, David G. Kirkpatrick, Jan Kratochvíl, and Igor Kríz. On restricted two-factors. *SIAM J. Discret. Math.*, 1(4):472–484, 1988.

Steven Jecmen, Hanrui Zhang, Ryan Liu, Nihar B. Shah, Vincent Conitzer, and Fei Fang. Mitigating manipulation in peer review via randomized reviewer assignments. In *Advances in Neural Information Processing Systems 33 (NeurIPS '20)*, 2020.

Ari Kobren, Barna Saha, and Andrew McCallum. Paper matching with local fairness constraints. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '19)*, pages 1247–1257. ACM, 2019.

K. Leyton-Brown and Mausam. Aaai 2021 - introduction. https://slideslive.com/38952457/aaai-2021-introduction?ref=account-folder-79533-folders, 2021. minute 8 onwards in the video.

Jing Wu Lian, Nicholas Mattei, Renee Noble, and Toby Walsh. The conference paper assignment problem: Using order weighted averages to assign indivisible goods. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI '18)*, pages 1138–1145. AAAI Press, 2018.

Michael L. Littman. Collusion rings threaten the integrity of computer science research. *Commun. ACM*, 64(6):43–44, 2021.

Cheng Long, Raymond Chi-Wing Wong, Yu Peng, and Liangliang Ye. On good and fair paper-reviewer assignment. In *Proceedings of the 2013 IEEE 13th International Conference on Data Mining (ICDM '13)*, pages 1145–1150. IEEE Computer Society, 2013.

Simon Price and Peter A. Flach. Computational support for academic peer review: a perspective from artificial intelligence. *Commun. ACM*, 60(3):70–79, 2017.

Nihar B. Shah. Systemic challenges and solutions on bias and unfairness in peer review. http://www.cs.cmu.edu/~nihars/preprints/SurveyPeerReview.pdf, 2021.

Ivan Stelmakh, Nihar Shah, and Aarti Singh. Peerreview4all: Fair and accurate reviewer assignment in peer review. *J. Mach. Learn. Res.*, 22(163):1–66, 2021.

Camillo J Taylor. On the optimal assignment of conference papers to reviewers. https://repository.upenn.edu/cis_reports/889, 2008.

Yichong Xu, Han Zhao, Xiaofei Shi, and Nihar B. Shah. On strategyproof conference peer review. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence (IJCAI '19)*, pages 616–622. ijcai.org, 2019.