Impartial Peer Review*

David Kurokawa[1], Omer Lev[2], Jamie Morgenstern[1], and Ariel D. Procaccia[1]

[1] *Computer Science Department, Carnegie Mellon University*
[2] *School of Computer Science and Engineering, The Hebrew University of Jerusalem*

Abstract

Motivated by a radically new peer review system currently under evaluation by the National Science Foundation, we study peer review systems in which proposals are reviewed by PIs who have submitted proposals themselves. An (m, k)-selection mechanism asks each PI to review m proposals, and uses these reviews to select (at most) k proposals. We are interested in *impartial* mechanisms, which guarantee that the ratings given by a PI to others' proposals do not affect the likelihood of the PI's own proposal being selected. We design two impartial mechanisms that select a k-subset of proposals that is essentially as highly rated as the one selected by the non-impartial (abstract version of) the NSF pilot mechanism, even when the latter mechanism has the "unfair" advantage of eliciting honest reviews.

*Kurokawa and Procaccia were partially supported by the NSF under grant CCF-1215883. Morgenstern was partially supported by NSF grants CCF-1116892 and CCF-1101215. This work has been partly supported by COST Action IC1205 on Computational Social Choice. Contact information: D. Kurokawa, J. Morgenstern, and A. D. Procaccia, Computer Science Department, Carnegie Mellon University, {dkurokaw,jamiemmt,arielpro}@cs.cmu.edu. O. Lev, School of Computer Science and Engineering, The Hebrew University of Jerusalem, omerl@cs.huji.ac.il.

1 Introduction

The Sensors and Sensing Systems (SSS) program of the National Science Foundation (NSF) is currently evaluating a revolutionary peer review method. Traditionally, grant proposals submitted to a specific program are evaluated by a panel of reviewers. Potential conflicts of interest play a crucial role in composing the panel; most importantly, principal investigators (PIs) whose proposal is being evaluated by the panel cannot serve on the panel. In stark contrast, the new peer review method — originally designed by Merrifield and Saari [11] for the review of proposals for telescope time — requires the PIs themselves to review each other's proposals! A "dear colleague letter" [9] explains the potential merits of the new process:

> *"This pilot is an attempt to find an alternative proposal review process that can preserve the ability of investigators to submit multiple proposals at more than one opportunity per year while encouraging high quality and collaborative research, placing the burden of proposal review onto the reviewer community in proportion to the burden each individual imposes on the system, simplifying the internal NSF review process, ameliorating concerns of conflict-of-interest, maintaining high quality in the review process, and substantially reducing proposal review costs."*

Under the Saari-Merrifield mechanism, each PI must review m proposals submitted by other PIs; in the NSF pilot, $m = 7$. The PI then ranks the m proposals according to their quality. To aggregate these rankings, the Borda count voting rule is used, so, essentially, each PI awards $m - i$ points to the proposal she ranks in position i. A proposal's overall rating is the average over the points awarded by the m PIs who reviewed it. An intriguing innovation is that a PI's own proposal receives a small bonus based on the similarity between the PI's submitted ranking and the aggregate ranking of proposals; this is meant to encourage PIs to make an effort to produce accurate reviews.

The NSF pilot sparked a lively debate among researchers in mechanism design and social choice, which also took place in the blogosphere [14, 16, 12]. While most researchers seem to agree that the NSF should be commended for trying out an ambitious peer review method, serious concerns have been raised regarding the pilot mechanism itself. Perhaps most strikingly, while the NSF announcement [9] states that the "theoretical basis for the proposed review process lies in an area of mathematics referred to as mechanism design", the pilot mechanism provides no theoretical guarantees. In particular, the mechanism is susceptible to strategic manipulation: PIs will often be able to advance their own proposals by giving low scores to competitive proposals (even though they may forfeit some of the small bonus for accurate reviewing). Indeed, while most researchers who sit on NSF panels are well-respected, the pilot mechanism cannot control the quality of PIs who submit proposals and therefore conduct reviews — leaving open the very real possibility of game-theoretic mayhem.

In this paper, we wish to rigorously study the design of peer review mechanisms where each reviewer is also associated with a proposal or a paper. These mechanisms must be *impartial*: reviewers must not be able to affect the chances of their own proposals being selected. Our research challenge is therefore to

> ... *design provably impartial peer review mechanisms that provide formal quality guarantees.*

We believe that solutions to this problem truly matter. The NSF plays a huge role in enabling scientific research in the United States, and its consideration of alternative peer review methods may transform how scientific funding is allocated in the US. The need to build sound foundations for these methods therefore provides a timely and unique opportunity for algorithmic game theory research.

1.1 Our Approach

In our setting there are n PIs, each associated with a proposal. Each PI i has a hypothetical (honest) evaluation of the quality of the proposal j, which is the rating i would give j if she were asked to review that proposal. The *score* of a proposal is the average rating given to it by other PIs. As NSF program directors, if our budget is sufficient to fund k proposals, we would ideally want to select a set of k proposals with maximum score. Thus we distill the strategic aspects of the NSF reviewing setting and abstract away some other practical aspects, such as the fact that PIs may submit multiple proposals to the same program. However, as we discuss in §3 and §6, our model and results easily extend. Even under the simplified formulation, there are two obstacles we must overcome: we cannot possibly ask each PI to review all other proposals, and the reviews may not be honest.

To address the first problem, we restrict our mechanisms to m reviews per PI (much like the NSF pilot). We therefore define an (m, k)*-selection mechanism* as follows. First, the mechanism asks each PI to review m proposals, in a way that each proposal is reviewed by exactly m PIs; for every such pair (i, j), PI i's evaluation for proposal j is revealed. Based on these elicited reviews, the mechanism selects k vertices. The most natural (m, k)-selection mechanism is an abstract version of the NSF pilot mechanism, which we fondly refer to as the Vanilla mechanism; it chooses m reviews per PI uniformly at random (subject to the constraint that each proposal is reviewed by m PIs), and then selects the k vertices with highest average rating, based only on the sampled reviews.

Returning to the second problem — dishonest reviewing — we say that a selection mechanism is *impartial* if the probability of proposal i being selected is independent of the ratings given by PI i. The motivation for our work stems from the observation that the Vanilla mechanism is not impartial: we seek mechanisms that are.

But how should we evaluate the impartial mechanisms we design? Without any assumptions, competing with an omniscient mechanism that maximizes underlying scores is clearly impossible. We therefore use the Vanilla mechanism as our performance benchmark. Competing with Vanilla is nontrivial, because we give it the "unfair" advantage of assuming that reviews are honest, even though it is not impartial. Specifically, we say that an impartial mechanism α-*approximates* Vanilla if, in the worst case over reviews, the ratio between the expected score (based on the largely unseen set of all possible reviews) of the set of proposals selected by the impartial mechanism, and the expected score of the set of proposals selected by Vanilla, is at least α.

The choice of Vanilla as a benchmark has two main advantages. First, since the Vanilla Mechanism is an abstraction of the NSF pilot mechanism, our choice of benchmark allows us to quantify how much the NSF must sacrifice to achieve impartiality — and our results show that this sacrifice is negligible. Moreover, our experience with the theory and practice of market design — specifically, kidney exchange (see, e.g., [2, 3, 4]) — has taught us that innovations that are closest to the current accepted practice are the most likely to be adopted.

Second, modulo its lack of impartiality, Vanilla is a good mechanism. In particular, in an average-case model where each proposal has an intrinsic quality, and reviews are drawn from a well-behaved distribution whose expectation is the true quality, Vanilla will pinpoint the best proposals given a sufficiently large m. Any approximation in our worst-case model carries over, of course, to this average-case model. And it makes sense to focus on the worst-case model because, as we discuss below, *even* in that model we can obtain excellent approximations of Vanilla via impartial mechanisms.

1.2 Our Results

We present two impartial (m, k)-selection mechanisms. In §3 we introduce and analyze the Partition mechanism, which approximates Vanilla to a factor of $(1 - \epsilon(k))$, where $\epsilon(k) = o(1)$, as long as m is significantly smaller than k. We think of m, the number of reviews per PI, as being a small constant, and we would like to think of k, the number of proposals to be selected, as significantly larger. (Here we are being somewhat optimistic about NSF funding! But see §6.)

We view Partition as practical, as it modifies Vanilla in a rather minimal way while guaranteeing impartiality. Indeed, it simply partitions the PIs into subsets, and only considers a review if the proposal is in one subset and its reviewer is in a different subset. We also informally argue that this mechanism can easily incorporate practical extensions.

Our second impartial mechanism, Credible Subset, presented in §4, (usually) selects k proposals at random from a slightly larger pool (of size $k + m$) of eligible, proposals. While it is not as practical as Partition, it gives a sharper approximation ratio of $\frac{k}{k+m}$. In §5, we show that the bound given by Credible Subset is, indeed, asymptotically tight when $k = m^2$ is a constant and the number of PIs n grows.

1.3 Related Work

Our paper is closely related to the work of Alon et al. [1]. In parallel with Holzman and Moulin [10], whose work we discuss below, Alon et al. introduced the notion of impartial selection mechanisms (using the term "strategyproofness" for impartiality). Their model can be interpreted as a special case of our model, where $m = n-1$ (i.e., each PI reviews all other proposals) and all the ratings are in $\{0, 1\}$. The main result of Alon et al. is the design of an impartial mechanism that approximates *the score of the optimal subset of k vertices* to a factor that goes to 1 as k grows. When $m = n - 1$ and all ratings are in $\{0, 1\}$, this is equivalent to approximating Vanilla: Vanilla can see all ratings and will select the optimal subset. For the case of $k = 1$, the 4-approximation result of Alon et al. was recently improved to a 2-approximation by Fischer and Klimm [7].

It is not hard to extend the results of Alon et al. [1] to ratings that are not in $\{0, 1\}$. However, designing mechanisms with approximation guarantees for all m in the range $\{1, \ldots, n-1\}$ makes our task much harder. Indeed, motivated by the idea that researchers can perform only a limited number of reviews, (m, k)-selection mechanisms are limited to only m reviews per PI, greatly reducing the amount of information available to the mechanism. In contrast, the k-selection mechanisms of Alon et al., which are mainly motivated by social networks, necessarily have complete information. This is crucial on a conceptual level, and motivates our use of Vanilla as a benchmark for our approximation results, rather than comparing directly to the optimal solution. It is also crucial on a technical level, as for $m \ll n-1$ we cannot reason about scores directly, as Alon et al. do. In fact, in this regime, which is typical for a peer review setting, our results are incomparable to theirs:

our mechanisms use far less information, but the performance of these mechanisms is (necessarily) measured against a weaker benchmark.

Independently of Alon et al. [1], Holzman and Moulin [10] introduced and studied impartial mechanisms in a similar model (in particular, mechanisms with complete information). They focused on the case of $k = 1$: only a single vertex is to be selected. Their approach is axiomatic: rather than optimizing an objective function subject to impartiality, they look for mechanisms that are impartial and satisfy a number of other normative properties.

Merrifield and Saari [11] are not the first researchers to suggest improvements to the peer review process, although most other papers focus on conference reviewing. For example, Douceur [6] proposes to use rankings instead of ratings for the evaluation of conference submissions. Roos et al. [15] propose a method for calibrating the ratings of potentially biased reviewers via a maximum likelihood estimation (MLE) approach. Similarly, Haenni [8] views reviewer scores as probabilities, and aggregates them using *Dempster's rule of combination*, a notion from the literature on reasoning under uncertainty. And Nierstrasz [13] presents a set of *process patterns*, which are used to categorize papers in a way that corresponds to how PC members are expected to behave in a PC meeting (e.g., the category "good paper" is interpreted as "I will champion the paper"). All these papers deal with the effective aggregation of reviewer opinions, whereas we deal with the (arguably) more basic problem of eliciting accurate opinions in the first place.

2 The Model

Let $N = \{1, 2, \ldots, n\}$ be the set of proposals and also the set of strategizing reviewers. Each reviewer i has an estimate of the quality of every other proposal $j \neq i$ — the score i would give j if i honestly reviewed j. We represent this setting as a weighted, complete, directed graph $G = (N, E, w_G)$ where $E = \{(i, j) \mid i, j \in N, i \neq j\}$, and $w_G(i, j) \in \mathbb{R}^+$ is the quality of j according to i's evaluation. We call G the *underlying graph*.

Let m be the number of proposals that each PI can review, which must equal to the number of reviews each proposal receives (we assume each PI submits one proposal, but see §3). In our model, m is the number of outgoing edges from each vertex and the number of incoming edges to each vertex. Slightly abusing terminology, we say that a directed graph is *m-regular* if it satisfies these properties.

A peer review process is governed by an *(m, k)-selection mechanism*, which works in two stages:

1. The mechanism selects (possibly randomly) a directed m-regular graph $G^m = (N, E(G^m))$, called the *sampled graph*. We assume this graph is drawn prior to the next step: that the sampling is done all at once independent of the edge weights.

2. Given the underlying graph G, the weight $w_G(i, j)$ is revealed for each edge $(i, j) \in E(G^m)$. The mechanism then maps these elicited ratings to a subset of selected vertices of size at most k.

Step 1 corresponds to the mechanism assigning m proposals to each PI. Based on the reviews $w_G(i, j)$ for $(i, j) \in E(G^m)$, in Step 2, the mechanism selects a subset of at most k proposals that will receive funding.

Let us reinterpret the NSF pilot mechanism [9] in this framework, abstracting away details such as the use of Borda count and the bonus component for accurate reviews. To this end, let $\mathcal{G}^m$

denote the uniform distribution over m-regular graphs. Given a weighted m-regular graph G^m, let

$$\mathrm{top}_k(G^m) \in \operatorname*{arg\,max}_{Y \subseteq N:\ |Y|=k} \sum_{i \in Y} \sum_{j:(j,i) \in E(G^m)} w_G(j,i),$$

breaking ties lexicographically. Now, the *Vanilla* mechanism, denoted $\mathcal{M}^v$, is defined as follows:

<u>Vanilla</u> (G, m, k)

1. Draw $G^m \sim \mathcal{G}^m$.
2. Return $\mathrm{top}_k(G^m)$.

Intuitively, the mechanism assigns proposals to PIs for review based on the graph G^m, and then returns the k highest-rated reviews based on the sampled reviews (for convenience we look at the sum of ratings, which is equivalent to the average).

For a mechanism $\mathcal{M}$ and an underlying graph G, let $\mathcal{M}(G)$ be a random variable, which takes the value $X \subseteq N$ with the same probability that $\mathcal{M}$ outputs X when the underlying graph is G. Then we can use $\mathbb{P}[i \in \mathcal{M}(G)]$ to denote the probability that $\mathcal{M}$ selects $i \in N$ when the underlying graph is G. We say that $\mathcal{M}$ is *impartial* if for any $i \in N$ and any two underlying graphs G and G' that differ only in the weights on the outgoing edges of i, $\mathbb{P}[i \in \mathcal{M}(G)] = \mathbb{P}[i \in \mathcal{M}(G')]$.

Unfortunately, Vanilla is blatantly not impartial. To see this, let $k = 1$, $m = 1$, and define the weights of G and G' as follows:

$$w_G(i,j) = \begin{cases} n+1 & i = 1 \\ 1 & j = 1 \\ 0 & \text{otherwise} \end{cases} \qquad w_{G'}(i,j) = \begin{cases} 0 & i = 1 \\ 1 & j = 1 \\ 0 & \text{otherwise} \end{cases}$$

Then $\mathbb{P}[1 \in \mathcal{M}^v(G)] = 0$, whereas $\mathbb{P}[1 \in \mathcal{M}^v(G')] = 1$ (using lexicographic tie-breaking, 1 would be selected even if only 0-weight edges are sampled).

The purpose of this paper is to design (m,k)-selection mechanisms that are simultaneously impartial (unlike Vanilla), yet similarly practical in terms of the number of reviews per proposal and similar in the quality of the output. We measure the quality of a mechanism by the expected score of the vertices it selects. Formally, let $\mathrm{sc}(i,G) = \sum_{(j,i) \in E} w_G(j,i)$ be the *score of node i in G*, and let $\mathrm{sc}(X,G) = \sum_{i \in X} \mathrm{sc}(i,G)$ be the score of a set of vertices $X \subseteq N$ in G. We can now define

$$\mathrm{sc}(\mathcal{M}, G) = \mathbb{E}_{X \sim \mathcal{M}(G)}[\mathrm{sc}(X,G)].$$

This is our optimization objective.

Note that for some underlying graphs G, Vanilla itself may do poorly in terms of $\mathrm{sc}(\mathcal{M}, G)$. As an extreme example, let $k = 1$, $m = 1$, and define the weights of the underlying graph G as follows:

$$w_G(i,j) = \begin{cases} 1000 & i = 1 \wedge j = 2 \\ 1/n & j = 1 \\ 0 & \text{otherwise} \end{cases}$$

It is very likely that the edge $(1,2)$ will not be sampled by Vanilla, and therefore the mechanism will likely select vertex 1. However, $\mathrm{sc}(1,G) = \frac{n-1}{n} < 1$, whereas $\mathrm{sc}(2,G) = 1000$. This is not a shortcoming of VANILLA specifically — it is clear that such examples can be constructed for any (m, k)-selection mechanism when m is much smaller than n.

Nevertheless, we can use VANILLA as a benchmark. We wish to design impartial mechanisms whose quality guarantee is quite close to that of VANILLA pointwise (assuming all reviews given to VANILLA were truthful). We say that an (m, k)-selection mechanism $\mathcal{M}$ α-approximates VANILLA, for $\alpha = \alpha(m, n, k) \leq 1$, if for *every* underlying graph G,

$$\frac{\mathrm{sc}(\mathcal{M}, G)}{\mathrm{sc}(\mathcal{M}^v, G)} \geq \alpha.$$

Before turning to the mechanism design task at hand, let us illustrate a technical subtlety of our approach through an example. Let $k = 2$, $m = O(1)$, and define the weights of the underlying graph G as follows:

$$w_G(i, j) = \begin{cases} 2^n & i = 1, j = 2 \\ 1 & i \neq 1, j = 2 \\ 2 & j = 1 \\ 0 & \text{otherwise} \end{cases}$$

Regardless of which edges are sampled, $\mathrm{top}_k(G^m) = \{1, 2\}$. But regardless of the sampling scheme[1], it is unlikely that the heavy edge $(1, 2)$ will be sampled. Then, based on the sampled graph, vertex 1 almost always has a higher sampled score than 2. However, most of $\mathrm{sc}(\{1, 2\}, G)$ comes from from vertex 2. VANILLA always selects $\{1, 2\}$: thus, in order to achieve worst-case guarantees with respect to VANILLA, it is not enough to select vertices that VANILLA deems more valuable. In fact, a mechanism must closely imitate VANILLA's entire distribution over selected vertices to achieve a good approximation to VANILLA. In other words, every vertex selected by VANILLA with some probability must be selected by our mechanisms with similar probability. We view this as a desirable feature of the worst-case model: as noted above, new mechanisms are more likely to be adopted when they are similar to established mechanisms, and we are able to design *impartial* mechanisms that closely imitate VANILLA in the foregoing sense.

3 The Partition Mechanism

We now present the *Partition* Mechanism, an impartial (m, k)-selection mechanism. In the formal description of the mechanism below, we denote it by $\mathcal{M}^p$, and use the constants a and b, which will be discussed later on.

PARTITION(G, m, k)

1. Draw $G^m \sim \mathcal{G}^m$.
2. $S_1, \ldots, S_{k^a} \leftarrow \emptyset$.
3. For all $i \in N$, assign i to one of $S_1, \ldots, S_{k^a}$ uniformly at random.
4. $H^m \leftarrow (N, \{(i, j) \in E(G^m) : \; i \in S_t, j \in S_{t'}, t \neq t'\})$.

[1] Recall that the graph is sampled independently of review scores, so there will always be some edge which is sampled with probability at most m/n.

5. For all $t \in [1, \ldots, k^a]$:

   (a) $\hat{X}_t \leftarrow \text{top}_{\frac{k}{k^a}+k^b}(H^m)$.

   (b) $X_t \leftarrow$ random subset of $\hat{X}_t$ of size k/k^a

6. Output $\bigcup_{t=1}^{k^a} X_t$.

We are assuming in the mechanism's description (and in our proofs) that all the numbers (such as k/k^a) are integers purely for ease of exposition. The intuition behind the mechanism is as follows. First, sample $G^m \sim \mathcal{G}^m$. Then, partition the vertices into k^a subsets, uniformly at random. Then, only request reviews from those edges $(i, j) \in G^m$ where i and j reside in distinct subsets. For each subset, choose the top $k/k^a + k^b$ vertices based only on these reviews, and then choose at random a subset of size k/k^a to output.

Why is Partition Impartial? After pruning G^m to get H^m, each vertex only rates vertices in other subsets. In particular, the selection of the top vertices of the subset that contains vertex i is only determined by edges from other subsets, so i cannot affect its own selection. This is true for any fixed graph G^m, which means that the probability of selecting i is independent of the weights on its outgoing edges.

Let us discuss several important points pertaining to Partition. First, the framework given by Partition is practically appealing, because it provides a natural way of dealing with conflicts of interest, by deliberately placing PIs who have a conflict in the same subset. As a special case, a PI who submits multiple proposals can be represented as multiple vertices that reside in the same subset. While, for ease of exposition, our model does not explicitly incorporate this practical extension, the analysis of Partition goes through almost unchanged.

Second, a naïve attempt at designing Partition might not include the step which picks uniformly at random from the top $k/k^a + k^b$ vertices in each subset, and instead choose the top k/k^a vertices from each subset. However, this mechanism might have arbitrarily bad approximation to Vanilla. This can be shown by constructing an example that builds on the idea presented at the end of §2: one of the vertices that is ranked lower by Vanilla can actually be the one yielding much of the score, if this is vertex has a high-weight incoming edge that is unlikely to be sampled. Partition therefore expands its search, in a sense, to make sure that every vertex selected by Vanilla is likely to be selected by Partition.

Third, Partition can be viewed as a direct extension of a similar mechanism introduced by Alon et al. [1]. Both mechanisms are strategyproof by the same (trivial) argument. But, for the reasons discussed in §1.3, the approximation analysis for our mechanism is quite different and more intricate.

Without further ado, we present our main technical result.

Theorem 1. Partition *is an impartial* (m, k)*-selection mechanism. Moreover, if* $1 - a > b > 1/2$ *and* $1/2 > a > 0$*, then* Partition $(1-\epsilon(k))$*-approximates* Vanilla*, for* $\epsilon(k) = o(1)$ *and* $m = o(k^a)$.

The theorem's formal proof is relegated to Appendix A. We have already argued that the mechanism is impartial, so we just need to establish its approximation guarantees. Informally, the analysis relies on two high-level ideas: (i) the vertices selected by Vanilla are almost evenly distributed between the subsets $S_1, \ldots, S_{k^a}$; (ii) fixing a sampled graph G^m, each vertex selected by Vanilla is unlikely to lose any edges when G^m is pruned to obtain H^m, so each such vertex is still likely to be selected. While establishing (i) is straightforward as the vertices are partitioned i.i.d., (ii) requires a fair amount of work to handle dependencies.

Note that the theorem assumes an upper bound on m; we reiterate that it is very natural to assume that it is small (constant, even), because each PI can only handle so many reviews.

4 The Credible Subset Mechanism

In this section we introduce our second mechanism, the *Credible Subset* mechanism. While it is not nearly as conceptually appealing as PARTITION— in particular, it cannot naturally handle conflicts of interest, and it makes some fairly arbitrary choices — it will allow us to explore the technical limits of our worst-case model, by providing a tight approximation ratio to VANILLA.

The mechanism relies on two ideas:

1. Every vertex that has the potential to be among the top k by changing its outgoing edges has a chance to be selected. Such vertices are called *credible.* There are not so many of them, and they include the actual top k.
2. A credible vertex can potentially affect the number of credible vertices, and therefore the probability of selecting a credible vertex must be independent of the number of credible vertices.

The Credible Subset mechanism, denoted $\mathcal{M}^{cs}$, formally works as follows.

CREDIBLE SUBSET(G, m, k)

1. Draw $G^m \sim \mathcal{G}^m$.
2. $P \leftarrow \{i \notin \text{top}_k(G^m) \mid \text{if } i \text{ reported } \forall j : w(i,j) = 0,\ i \text{ would be in top}_k(G^m)\}$
3. $S \leftarrow \text{top}_k(G^m) \cup P$.
4. With probability $\frac{|S|}{k+m}$ return a random k-subset of S, and with probability $1 - \frac{|S|}{k+m}$ return $\emptyset$.

Theorem 2. CREDIBLE SUBSET *is an impartial* (m, k)*-selection mechanism which* $\frac{k}{k+m}$*-approximates* VANILLA.

The theorem's simple proof (which just formalizes the two ideas given above) is relegated to Appendix B. In contrast to PARTITION, the (somewhat) harder part of the analyis is showing that CREDIBLE SUBSET is impartial — bounding its approximation ratio is trivial. It is important to note that when $m = o(k)$, the approximation ratio goes to 1 as k goes to infinity, similarly to Theorem 1 for PARTITION; but the guarantee given here is crisper.

We also remark that the mechanism may return subsets of size smaller than k — empty subsets, in fact! Choosing empty subsets is not necessary: the same approximation guarantee can be achieved by defining a finer distribution over subsets preserving that each vertex in S is selected with probability given by Lemma 6. Since this particular mechanism is mostly of theoretical interest, though, we prefer to focus on the simpler formulation of the mechanism for ease of exposition.

5 Impossibility Results

In §4 we proved that Credible Subset approximates Vanilla to a factor of $\frac{k}{k+m}$. When $m = o(k)$, this is $1 - o(1)$. But when both k and m are constants, this ratio is bounded away from 1 even when $n \to \infty$. It is natural to wonder, though, if an impartial (m, k)-selection mechanism can approximate Vanilla to a factor of $1 - o(1)$ when k and m are constants and n grows. After all, in this regime the performance of Vanilla will be very poor in the worst case (as G^m gives an extremely incomplete picture of G), so Vanilla becomes easier to approximate. We answer this question in the negative: we show below that the $\frac{k}{k+m}$ ratio is essentially the best possible for impartial (m, k)-selection mechanisms.

Let us start with an informal discussion of a simple upper bound of $\frac{k}{k+1}$ that only assumes that $k \le m$ (that is, it gives a constant upper bound for $k = O(1)$ even if m grows). Let G be an underlying graph such that

$$w_G(i,j) = \begin{cases} \epsilon & j = 1 \\ 0 & \text{otherwise} \end{cases}$$

Vanilla will certainly select vertex 1. Consider an impartial (m, k)-selection mechanism $\mathcal{M}$, and let $\mathbb{P}[1 \in \mathcal{M}(G)] = p$. Since 1 is the only vertex with nonzero score, the approximation ratio of $\mathcal{M}$ on G is p.

Next, consider the underlying graph G' with weights defined by

$$w_{G'}(i,j) = \begin{cases} \epsilon & j = 1 \\ 1 & i = 1 \\ 0 & \text{otherwise} \end{cases}$$

For $\epsilon \ll \frac{1}{n-1}$, Vanilla will certainly select k vertices with score 1, so $\mathrm{sc}(\mathcal{M}^v, G') = k$. By impartiality, $\mathbb{P}[1 \in \mathcal{M}(G')] = p$, hence

$$\mathrm{sc}(\mathcal{M}, G') \le (1-p)k + p(k - 1 + (n-1)\epsilon).$$

Since ϵ is arbitrarily small, the approximation ratio is upper-bounded in the limit by

$$\alpha = \min\left\{p, (1-p) + \frac{p(k-1)}{k}\right\}.$$

Maximizing α over all $p \in [0,1]$ gives $p = \frac{k}{k+1}$ as an upper bound on the approximation ratio.

Let us now turn to our more intricate upper bound.

Theorem 3. *Let $c \in (0, \frac{1}{4})$, $k = m^2$, and $m \le n^c$. Then any impartial mechanism at best $\left(\frac{k}{k+m} + \epsilon(n)\right)$-approximates* Vanilla, *for $\epsilon(n) = o(1)$.*

The formal proof appears in Appendix C. In a nutshell, we construct an underlying graph with a set of vertices that have low (but nonzero) score. Vanilla selects k valuable vertices outside this set. To compete, an impartial mechanism must select many of the same vertices. But in a different graph, one of the vertices in that set has an extremely valuable incoming edge. Vanilla is unlikely to see this edge, and so is a given impartial (m, k)-selection mechanism (in fact, for any given mechanism, we specifically choose this edge so it is unlikely to be sampled). But we construct a graph so that Vanilla is lucky enough to select the high-score vertex, even though the high-weight edge is not sampled; whereas any impartial mechanism is not very likely to select the high-score vertex, due to its performance on the first graph.

We remark that Alon et al. [1] prove an upper bound of $\frac{k^2+k-1}{k^2+k}$ for their setting, which is the special case of ours in the regime $m = n - 1$. They do this by creating a graph where all edges have weight 0 except for a cycle of length $k + 1$ of edges of weight 1. One of the vertices in this cycle — call it i — is selected with probability at most $k/(k+1)$. The upper bound is obtained by reducing the weight on i's outgoing edge to 0. In this new graph, i is still selected with probability at most $\frac{k}{k+1}$ by impartiality, so the mechanism's score is at most $\frac{k}{k+1}k + \frac{1}{k+1}(k-1)$, whereas the optimal solution (which is equivalent to Vanilla in this regime) achieves score k. It is interesting to note that this argument does not extend to the case of $m \ll n$, because Vanilla is unlikely to see the cycle of valuable edges.

6 Discussion

Application to Conference Reviewing. While we have focused on NSF reviewing in the introduction (and, indeed, this is the real-world setting that motivated us), our results can certainly be applied to conference reviewing. For example, in large conferences such as AAAI and IJCAI, the PC includes hundreds of people — a large fraction of the researchers who actually submit papers to the conference. These conferences are a great fit with our model and results, because: (i) Vanilla is, essentially, the mechanism that is typically used (modulo choosing the m-regular graph in a way that matches reviewers with suitable papers), and (ii) k (the number of papers selected for presentation and publication) is much larger than m (the number of reviews per PC member) — in IJCAI'13, the values were $k = 413$ and $m < 10$, making the Partition Mechanism (or a variation thereof) eminently practical.

Rankings and Borda scores. One of the ways in which the mechanism of Merrifield and Saari [11] differs from our setting is that reviewers are restricted to ranking the proposals. Since Borda count is used to aggregate the rankings, this is equivalent to limiting the reviewers to handing out the ratings $m-1, m-2, \ldots, 0$ (exactly one of each) — even though their true ratings may be different. But our mechanisms readily extend to this setting. Intuitively, this is because our arguments revolve around the idea of acting similarly to Vanilla (despite the impartiality constraint), and Vanilla is similarly affected. The only subtlety is that under one interpretation of Partition, PIs do not actually carry out the reviews in $G^m \setminus H^m$. If a PI ranks, say, only $m - 1$ proposals, the Borda scores assigned by the PI would be $m-2, m-3, \ldots, 0$, so discarding one edge may affect as many as $m - 1$ other scores. However, since very few edges are discarded, the analysis goes through.

"Universal impartiality" vs. "impartiality in expectation". The algorithmic mechanism design literature makes a distinction between randomized mechanisms that are *universally truthful* — players cannot gain from misreporting their private information regardless of the coin tosses of the mechanism — and *truthful in expectation* — players cannot improve their expected utility by misreporting, where the expectation is taken over the mechanism's coin tosses (see, e.g., [5]). Our definition of impartiality is *in expectation*, but both our mechanisms are actually *universally impartial*. Hence, our positive results are a bit stronger than stated, while our negative results work *even* when one asks only for impartiality in expectation.

References

[1] N. Alon, F. Fischer, A. D. Procaccia, and M. Tennenholtz. Sum of us: Strategyproof selection from the selectors. In *Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge (TARK)*, pages 101–110, 2011.

[2] I. Ashlagi, F. Fischer, I. Kash, and A. D. Procaccia. Mix and Match: A strategyproof mechanism for multi-hospital kidney exchange. *Games and Economic Behavior*, 2014. Forthcoming; preliminary version in EC'10.

[3] A. Blum, A. Gupta, A. D. Procaccia, and A. Sharma. Harnessing the power of two crossmatches. In *Proceedings of the 14th ACM Conference on Electronic Commerce (EC)*, pages 123–140, 2013.

[4] J. P. Dickerson, A. D. Procaccia, and T. Sandholm. Failure-aware kidney exchange. In *Proceedings of the 14th ACM Conference on Electronic Commerce (EC)*, pages 323–340, 2013.

[5] S. Dobzinski and S. Dughmi. On the power of randomization in algorithmic mechanism design. *SIAM Journal on Computing*, 42(6):2287–2304, 2013.

[6] J. Douceur. Paper rating vs. paper ranking. *Operating Systems Review*, 43:117–121, 2009.

[7] F. Fischer and M. Klimm. Optimal impartial selection. In *Proceedings of the 15th ACM Conference on Economics and Computation (EC)*, 2014. Forthcoming.

[8] R. Haenni. Aggregating referee scores: an algebraic approach. In *Proceedings of the 2nd International Workshop on Computational Social Choice (COMSOC)*, pages 277–288, 2008.

[9] G. A. Hazelrigg. Dear colleague letter: Information to principal investigators (PIs) planning to submit proposals to the Sensors and Sensing Systems (SSS) program October 1, 2013, deadline. `http://www.nsf.gov/pubs/2013/nsf13096/nsf13096.jsp?WT.mc_id=USNSF_25#reference1`, 2013. Retrieved on June 17, 2014.

[10] R. Holzman and H. Moulin. Impartial nominations for a prize. *Econometrica*, 81(1):173–196, 2013.

[11] M. Merrifield and D. Saari. Telescope time without tears: a distributed approach to peer review. *Astronomy and Geophysics*, 50(4):2–6, 2009.

[12] M. Mitzenmacher. NSF reviewing trial run. `http://mybiasedcoin.blogspot.com/2013/06/nsf-reviewing-trial-run.html`, 2013. Retrieved on June 17, 2014.

[13] O. Nierstrasz. Identify the champion. In N. Harrison, B. Foote, and H. Rohnert, editors, *Pattern Languages of Program Design*, volume 4, pages 539–556. Addison-Wesley, 2000.

[14] Ariel D. Procaccia. NSF (actually) reviewing via social choice. `http://agtb.wordpress.com/2013/06/10/nsf-actually-reviewing-via-social-choice/`, 2013. Retrieved on June 17, 2014.

[15] M. Roos, J. Rothe, and B. Scheuermann. How to calibrate the scores of biased reviewers by quadratic programming. In *Proceedings of the 25th AAAI Conference on Artificial Intelligence (AAAI)*, pages 255–260, 2011.

[16] R. V. Vohra. A mechanism design approach to peer review. `http://theoryclass.wordpress.com/2013/06/06/a-mechanism-design-approach-to-peer-review/`, 2013. Retrieved on June 17, 2014.

A Proof of Theorem 1

The proof relies on four lemmas. The first three lemmas analyze the probability of different events occurring; we will show these events become very unlikely as k grows (all the lemmas make the same assumptions regarding a and b as the theorem). In particular, the first lemma deals with having many high-scoring vertices in a single set, specifically some S_t containing more than $\frac{k}{k^a} + k^b$ of the $top_k(G^m)$ vertices.

Lemma 1. *$\lim_{k\to\infty} \mathbb{P}\left[\exists t, |S_t \cap top_k(G^m)| > \frac{k}{k^a} + k^b\right] = 0.$*

Proof. Let $\tilde{X}_t$ be the random variable denoting the number of elements of $top_k(G^m)$ which are placed in the set S_t. We view this as a sum of k Bernoulli trials with a bias of $\frac{1}{k^a}$. By the linearity of expectation, $\mathbb{E}[\tilde{X}_t] = \frac{k}{k^a}$. Therefore, by Hoeffding's inequality, for $\epsilon \in [0,1]$, we have

$$\mathbb{P}\left[\tilde{X}_t > \frac{k}{k^a} + k\epsilon\right] \le e^{-2k\epsilon^2}.$$

If $\epsilon = k^{b-1} < 1$ (which holds for $b < 1 - a < 1$), a union bound over the k^a sets yields

$$\mathbb{P}\left[\exists t, |S_t \cap \mathrm{top}_k(G^m)| > \frac{k}{k^a} + k^b\right] \le \sum_{t=1}^{k^a} \mathbb{P}\left[|S_t \cap \mathrm{top}_k(G^m)| > \frac{k}{k^a} + k^b\right] \le k^a e^{-2k^{2b-1}}.$$

When $b > \frac{1}{2}$ it holds that $2b - 1 > 0$, and therefore

$$\lim_{k\to\infty} \frac{k^a}{e^{2k^{2b-1}}} = 0.$$

□

The next two lemmas deal with the probability that vertices have no edges removed from them in the creation of H^m, thus retaining the same score (and hence, rating) as in G^m, and are both used in the proof of Lemma 4. First, looking at k^c (for $c < a$) of the top k vertices selected by Vanilla, the probability that any of them has lost an edge in H^m approaches 0 as k grows.

Lemma 2. *For every sampled graph G^m and $c < a$,*

$$\lim_{k\to\infty} \mathbb{P}[\exists i \in top_{k^c}(G^m),\ \exists j \in N,\ (j,i) \in G^m \setminus H^m \mid G^m] = 0.$$

Proof. In the proof we do not explicitly condition on G^m to simplify notation. For a given vertex $i \in N$, the probability that i loses no edges is

$$\mathbb{P}[\forall j, (j,i) \in E(G^m) \Rightarrow (j,i) \in E(H^m)] = \left(1 - \frac{1}{k^a}\right)^m.$$

Then, the probability a particular i loses an edge is

$$\mathbb{P}[\exists j,\ (j,i) \in E(G^m) \setminus E(H^m)] = 1 - \left(1 - \frac{1}{k^a}\right)^m,$$

and using the union bound over k^c vertices,

$$\begin{aligned}\mathbb{P}[\exists i \in \mathrm{top}_{k^c}(G^m),\ \exists j \in N,\ (i,j) \in G^m \setminus H^m] &\leq k^c \left(1 - \left(1 - \frac{1}{k^a}\right)^m\right) \\ &\leq k^c \left(1 - \left(1 - \frac{1}{k^a}\right)\right) \\ &= k^{c-a}.\end{aligned} \tag{1}$$

But because $c < a$, the right hand side of Equation (1) goes to 0 as k goes to infinity. □

Examining the distribution of vertices that retained all their edges in Partition's H^m, we order the vertices in $\mathrm{top}_k(G^m)$ according to their score, and we define the sets $V_0 = \{1, \ldots, k^c\}$ and for $1 \leq i \leq \log \frac{k}{k^c}$ $(c < a < 1)$, $V_i = \{2^{i-1}k^c + 1, \ldots, 2^i k^c\}$ according to this order. We denote

$$\hat{V}_i = \{i \in V_i \mid \forall j \in N,\ (i,j) \in E(G^m) \Rightarrow (i,j) \in E(H^m)\},$$

that is, $\hat{V}_i$ is the set of vertices in V_i that have all their edges in H^m. The next lemma states that the probability that for some i, $|\hat{V}_i| < 2^i k^c (1 - \frac{1}{k^a})^m - (2^i k^c)^d$ (for $0 < d < 1$) approaches 0 as k grows.

Lemma 3. *For every sampled graph G^m and every $0 < d < 1$,*

$$\lim_{k \to \infty} \mathbb{P}\left[\exists i,\ |\hat{V}_i| < 2^{i-1}k^c \left(1 - \frac{1}{k^a}\right)^m - (2^{i-1}k^c)^d \;\middle|\; G^m\right] = 0.$$

Proof. In the proof we do not explicitly condition on G^m to simplify notation. Let $s_i = 2^{i-1}k^c$, and let $n_i = s_i(1 - \frac{1}{k^a})^m - s_i^d$. For a given i, the probability that fewer than n_i vertices in V_i have lost any edges is equal to the probability that at least $s_i - n_i$ vertices in V_i have lost at least one edge. Formally. denoting

$$\hat{U}_i = V_i \setminus \hat{V}_i = \{i \in V_i \mid \exists j \in N,\ (j,i) \in E(G^m) \setminus E(H^m)\},$$

we have that:

$$\mathbb{P}\left[|\hat{V}_i| < n_i\right] = \mathbb{P}\left[|\hat{U}_i| \geq s_i - n_i\right].$$

The probability that a vertex loses an edge is $1 - (1 - \frac{1}{k^a})^m$. Using the linearity of expectation, it holds that

$$\mathbb{E}\left[\left|\hat{U}_i\right|\right] = s_i \left(1 - \left(1 - \frac{1}{k^a}\right)^m\right).$$

By Markov's inequality:

$$\begin{aligned}\mathbb{P}\left[\left|\hat{U}_i\right| \geq s_i - n_i\right] &\leq \frac{s_i - s_i(1 - \frac{1}{k^a})^m}{s_i - n_i} = \frac{s_i - s_i(1 - \frac{1}{k^a})^m}{s_i - s_i(1 - \frac{1}{k^a})^m + s_i^d} \\ &= \frac{1 - (1 - \frac{1}{k^a})^m}{1 - (1 - \frac{1}{k^a})^m + s_i^{d-1}} \leq \frac{1 - (1 - \frac{1}{k^a})^m}{1 - (1 - \frac{1}{k^a})^m + \left(2^{\log(k^{1-c})}k^c\right)^{d-1}} \\ &= \frac{1 - (1 - \frac{1}{k^a})^m}{1 - (1 - \frac{1}{k^a})^m + k^{d-1}},\end{aligned}$$

where the fourth transition follows from the fact that $s_i < 2^{\log k k^c} k^c$ and $d < 1$. A union bound over all $i \in \{0, \ldots, \log(\frac{k}{k^c})\}$ gives

$$\mathbb{P}[\exists i,\ |\hat{V}_i| < n_i] \leq (\log\left(k^{1-c}\right)+1)\left(\frac{1-(1-\frac{1}{k^a})^m}{1-(1-\frac{1}{k^a})^m + k^{d-1}}\right) \leq 2\log\left(k^{1-c}\right)\left(\frac{1-(1-\frac{1}{k^a})^m}{1-(1-\frac{1}{k^a})^m + k^{d-1}}\right). \tag{2}$$

To establish the lemma it is sufficient to show that the right hand side of Equation (2) goes to 0 as $k \to \infty$.

$$\begin{aligned}
2\log\left(k^{1-c}\right)\left(\frac{1-(1-\frac{1}{k^a})^m}{1-(1-\frac{1}{k^a})^m + k^{d-1}}\right) &= 2(1-c)\log(k)\left(\frac{1}{1+\frac{k^{d-1}}{1-(1-\frac{1}{k^a})^m}}\right) \\
= 2(1-c)\left(\frac{\log(k)}{1+\frac{1}{k^{1-d}(1-(1-\frac{1}{k^a})^m)}}\right) &= (1-c)\left(\frac{e^{\log(\log(k))}}{1+\frac{1}{e^{(1-d)\log(k)}(1-(1-\frac{1}{k^a})^m)}}\right) \\
= 2(1-c)&\left(\frac{1}{e^{-\log(\log(k))}(1+\frac{1}{e^{(1-d)\log(k)}(1-(1-\frac{1}{k^a})^m)})}\right) \\
= 2(1-c)&\left(\frac{1}{e^{-\log(\log(k))}+\frac{e^{-\log(\log(k))}}{e^{(1-d)\log(k)}(1-(1-\frac{1}{k^a})^m)}}\right) \\
= 2(1-c)&\left(\frac{1}{e^{-\log(\log(k))}+\frac{1}{e^{\log(\log(k))}e^{(1-d)\log(k)}(1-(1-\frac{1}{k^a})^m)}}\right) \\
= 2(1-c)&\left(\frac{1}{e^{-\log(\log(k))}+\frac{1}{e^{\log(\log(k))+(1-d)\log(k)}(1-(1-\frac{1}{k^a})^m)}}\right)
\end{aligned}$$

Since $e^{-\log(\log(k))} \xrightarrow[k\to\infty]{} 1$, to prove our claim it suffices to prove

$$\lim_{k\to\infty} e^{\log(\log(k))+(1-d)\log(k)}(1-(1-\frac{1}{k^a})^m) = 0.$$

Considering $1-(1-\frac{1}{k^a})^m = 1-((1-\frac{1}{k^a})^{k^a})^{\frac{m}{k^a}}$, we realize:

$$\begin{aligned}
\lim_{k\to\infty} e^{\log(\log(k))}e^{(1-d)\log(k)}(1-(1-\frac{1}{k^a})^m) &= \lim_{k\to\infty} e^{\log(\log(k))+(1-d)\log(k)}(1-e^{-\frac{m}{k^a}}) \\
&= \lim_{k\to\infty} e^{\log(\log(k))+(1-d)\log(k)} - e^{\log(\log(k))+(1-d)\log(k)-\frac{m}{k^a}} = 0
\end{aligned}$$

As was needed to complete the proof. □

The final lemma shows that, as k grows, the ratio of the expected score of the vertices in $\text{top}_k(G^m)$ that have lost no edges in PARTITION's H^m, to the score of $\text{top}_k(G^m)$, approaches 1. This lemma is the main tool used in the proof of Theorem 1. Denote

$$\hat{V} = \{i \in \mathrm{top}_k(G^m) \mid \forall j \in N,\ (j,i) \in E(G^m) \Rightarrow (j,i) \in E(H^m)\} = \bigcup_{t=0}^{\log(\frac{k}{k^c})} \hat{V}_t.$$

Lemma 4. *For every sampled graph G^m,*

$$\lim_{k\to\infty} \frac{\mathbb{E}[\sum_{i\in\hat{V}} sc(i,G) \mid G^m]}{\sum_{i\in top_k(G^m)} sc(i,G)} = 1.$$

Proof. In the proof we do not explicitly condition on G^m to simplify notation. For $1 \le i \le \log(\frac{k}{k^c})$, we define the set $\tilde{V}_i = \hat{V}_{i-1} \cup (V_i \setminus \hat{V}_i)$. Then

$$\bigcup_{i=1}^{\log(\frac{k}{k^c})} \tilde{V}_i = \mathrm{top}_k(G^m) \setminus \left((V_0 \setminus \hat{V}_-) \cup \hat{V}_{\log(\frac{k}{k^c})}\right).$$

Thanks to Lemma 2, $V_0 \setminus \hat{V}_0 = \emptyset$ with probability 1 as $k \to \infty$. Moreover, $\hat{V}_{\log(\frac{k}{k^c})}$ contains vertices for which no edge was removed in H^m. We also note that $\hat{V}_{i-1}$ is the set of vertices of $\tilde{V}_i$ from which no edge has been removed in H^m; formally,

$$\hat{V}_{i-1} = \{i \in \tilde{V}_i :\ \forall j \in N,\ (j,i) \in E(G^m) \Rightarrow (j,i) \in E(H^m)\}.$$

It follows that for any G^m, when $V_0 \setminus \hat{V}_0 = \emptyset$, if $\min_i \frac{\sum_{j\in\hat{V}_{i-1}} \mathrm{sc}(j,G)}{\sum_{j\in\tilde{V}_i} \mathrm{sc}(j,G)} = \ell$, then

$$\sum_{j\in\hat{V}} \mathrm{sc}(j,G) = \sum_i \sum_{j\in\hat{V}_{i-1}} \mathrm{sc}(j,G) \ge \ell \sum_i \sum_{j\in\tilde{V}_i} \mathrm{sc}(j,G).$$

And since $\sum_{j\in\mathrm{top}_k(G^m)} \mathrm{sc}(j,G) = \sum_i \sum_{j\in\tilde{V}_i} \mathrm{sc}(j,G) + \sum_{j\in\hat{V}_{\log(\frac{k}{k^c})}} \mathrm{sc}(j,G)$, we have:

$$\lim_{k\to\infty} \mathbb{P}\left[\frac{\sum_{j\in\hat{V}} \mathrm{sc}(j,G)}{\sum_{j\in\mathrm{top}_k(G^m)} \mathrm{sc}(j,G)} \ge \min_i \frac{\sum_{j\in\hat{V}_{i-1}} \mathrm{sc}(j,G)}{\sum_{j\in\tilde{V}_i} \mathrm{sc}(j,G)}\right] = 1.$$

Since in every $\tilde{V}_i$ the highest scoring elements are members of $\hat{V}_{i-1}$ (as the vertices were ordered according to score), we can ignore the scores, and just count the number of elements. Furthermore, Lemma 3 provides a bound on the size of $\hat{V}_i$ that holds with high probability. Now, using the notation from the proof of Lemma 3, where $s_i = 2^{i-1}k^c$ and $n_i = s_i(1-\frac{1}{k^a})^m - (s_i)^d$, for a fixed i, it holds with high probability that

$$\frac{\sum_{j\in\hat{V}_{i-1}} \mathrm{sc}(j,G)}{\sum_{j\in\tilde{V}_i} \mathrm{sc}(j,G)} \ge \frac{|\hat{V}_{i-1}|}{|\hat{V}_{i-1}| + |V_i| - |\hat{V}_i|} \ge \frac{|\hat{V}_{i-1}|}{|\hat{V}_{i-1}| + s_i - n_i} \ge \frac{n_{i-1}}{n_{i-1} + 2^i k^c - n_i}.$$

Now, substituting n_i and n_{i-1}, we get

$$\begin{aligned}
\frac{n_{i-1}}{n_{i-1} + s_i - n_i} &= \frac{s_{i-1}(1-\frac{1}{k^a})^m - (s_{i-1})^d}{s_{i-1}(1-\frac{1}{k^a})^m - (s_{i-1})^d + s_i - (s_i(1-\frac{1}{k^a})^m - (s_i)^d)} \\
&= \frac{(1-\frac{1}{k^a})^m - (s_{i-1})^{d-1}}{(1-\frac{1}{k^a})^m - (s_{i-1})^{d-1} + 2 - 2(1-\frac{1}{k^a})^m + 2^d(s_{i-1})^{d-1}} \\
&= \frac{(1-\frac{1}{k^a})^m - (s_{i-1})^{d-1}}{2 - (1-\frac{1}{k^a})^m + (2^d-1)s_{i-1}^{d-1}}.
\end{aligned} \tag{3}$$

It remains to give a lower bound on the right hand side of Equation (3), which is monotonically increasing with i because $d < 1$. Hence, we substitute s_{i-1} with its minimal value, which is k^c (for s_0), and use the assumption that $d < 1$, along with the theorem's assumption that $m = o(k^a)$:

$$\frac{(1-\frac{1}{k^a})^m - (s_{i-1})^{d-1}}{2-(1-\frac{1}{k^a})^m + (2^d-1)(s_{i-1})^{d-1}} \geq \frac{(1-\frac{1}{k^a})^m - k^{c(d-1)}}{2-(1-\frac{1}{k^a})^m + (2^d-1)k^{c(d-1)}} \xrightarrow{k\to\infty} 1$$

□

Using these lemmas, we can proceed to the theorem's proof.

Proof of Theorem 1. It holds that

$$\begin{aligned}\frac{\mathrm{sc}(\mathcal{M}^p, G)}{\mathrm{sc}(\mathcal{M}^v, G)} &= \frac{\sum_{G^m} \mathbb{P}[G^m](\sum_{i\in N} \mathrm{sc}(i,G)\cdot \mathbb{P}\left[i \in \mathcal{M}^p(G) \mid G^m\right])}{\sum_{G^m} \mathbb{P}[G^m] \sum_{i\in \mathrm{top}_k(G^m)} \mathrm{sc}(i,G)} \\ &\geq \frac{\sum_{G^m} \mathbb{P}[G^m](\sum_{i\in \mathrm{top}_k(G^m)} \mathrm{sc}(i,G)\cdot \mathbb{P}\left[i \in \mathcal{M}^p(G) \mid G^m\right])}{\sum_{G^m} \mathbb{P}[G^m] \sum_{i\in \mathrm{top}_k(G^m)} \mathrm{sc}(i,G)}\end{aligned}$$

To prove the theorem it is therefore sufficient to show that for every G^m,

$$\lim_{k\to\infty} \frac{\sum_{i\in \mathrm{top}_k(G^m)} \mathrm{sc}(i,G)\cdot \mathbb{P}\left[i \in \mathcal{M}^p(G) \mid G^m\right]}{\sum_{i\in \mathrm{top}_k(G^m)} \mathrm{sc}(i,G)} = 1. \tag{4}$$

To establish Equation (4), we fix a sampled graph G^m, but below we do not explicitly condition on G^m to simplify notation.

Since for every two events A and B it holds that $\mathbb{P}[A \mid B]\cdot\mathbb{P}[B] \leq \mathbb{P}[A]$, we have

$$\begin{aligned}&\sum_{i\in \mathrm{top}_k(G^m)} \mathrm{sc}(i,G)\cdot \mathbb{P}\left[i \in \mathcal{M}^p(G)\right] \\ &\quad\geq \sum_{i\in \mathrm{top}_k(G^m)} \mathrm{sc}(i,G)\cdot \mathbb{P}\left[i \in \mathcal{M}^p(G) \;\middle|\; i \in \hat{V} \wedge \forall t, |S_t \cap \mathrm{top}_k(G^m)| \leq \frac{k}{k^a} + k^b\right] \\ &\qquad\quad \cdot \mathbb{P}\left[i \in \hat{V} \wedge \forall t, |S_t \cap \mathrm{top}_k(G^m)| \leq \frac{k}{k^a} + k^b\right].\end{aligned}$$

Now suppose that indeed for all t, $|S_t \cap \mathrm{top}_k(G^m)| \leq \frac{k}{k^a} + k^b$. For any $1 \leq i \leq \log(\frac{k}{k^c})$, for any $j \in \hat{V}_i$ (i.e., j has not lost any edges during the construction of H^m), it must also hold that $j \in \mathrm{top}_k(H^m)$, and therefore it is in the $\mathrm{top}_{\frac{k}{k^a}+k^b}(S_t)$ of its respective group S_t. It follows that j has probability at least $\frac{\frac{k}{k^a}}{\frac{k}{k^a}+k^b}$ of being selected. Hence,

$$\begin{aligned}&\sum_{i\in \mathrm{top}_k(G^m)} \mathrm{sc}(i,G)\cdot \mathbb{P}\left[i \in \mathcal{M}^p(G) \;\middle|\; i \in \hat{V} \wedge \forall t, |S_t \cap \mathrm{top}_k(G^m)| \leq \frac{k}{k^a} + k^b\right] \\ &\qquad \cdot \mathbb{P}\left[i \in \hat{V} \wedge \forall t, |S_t \cap \mathrm{top}_k(G^m)| \leq \frac{k}{k^a} + k^b\right] \\ &\quad\geq \frac{\frac{k}{k^a}}{\frac{k}{k^a}+k^b} \sum_{i\in \mathrm{top}_k(G^m)} \mathrm{sc}(i,G)\cdot \mathbb{P}\left[i \in \hat{V} \wedge \forall t, |S_t \cap \mathrm{top}_k(G^m)| \leq \frac{k}{k^a} + k^b\right].\end{aligned} \tag{5}$$

Using $\mathbb{P}[A \wedge B] = \mathbb{P}[A] \cdot \mathbb{P}[B \mid A]$, we know that for $i \in \mathrm{top}_k(G^m)$,

$$\mathbb{P}\left[i \in \hat{V} \wedge \forall t, |S_t \cap \mathrm{top}_k(G^m)| \le \frac{k}{k^a} + k^b\right] = \mathbb{P}[i \in \hat{V}] \cdot \mathbb{P}\left[\forall t, |S_t \cap \mathrm{top}_k(G^m)| \le \frac{k}{k^a} + k^b \;\middle|\; i \in \hat{V}\right]. \tag{6}$$

We wish to examine $\mathbb{P}\left[\forall t, |S_t \cap \mathrm{top}_k(G^m)| \le \frac{k}{k^a} + k^b \;\middle|\; i \in \hat{V}\right]$, via

$$\begin{aligned}\mathbb{P}\left[\forall t, |S_t \cap \mathrm{top}_k(G^m)| \le \frac{k}{k^a} + k^b\right] &= \mathbb{P}\left[\forall t, |S_t \cap \mathrm{top}_k(G^m)| \le \frac{k}{k^a} + k^b \;\middle|\; i \in \hat{V}\right] \cdot \mathbb{P}\left[i \in \hat{V}\right] \\ &\quad + \mathbb{P}\left[\forall t, |S_t \cap \mathrm{top}_k(G^m)| \le \frac{k}{k^a} + k^b \;\middle|\; i \notin \hat{V}\right] \cdot \mathbb{P}\left[i \notin \hat{V}\right].\end{aligned}$$

As $\mathbb{P}\left[i \notin \hat{V}\right] \le \frac{m}{k^a}$ by the union bound, we know that

$$\mathbb{P}\left[\forall t, |S_t \cap \mathrm{top}_k(G^m)| \le \frac{k}{k^a} + k^b\right] \le \mathbb{P}\left[\forall t, |S_t \cap \mathrm{top}_k(G^m)| \le \frac{k}{k^a} + k^b \mid i \in \hat{V}\right] + \frac{m}{k^a}.$$

Plugging this into Equation (6),

$$\mathbb{P}\left[i \in \hat{V} \wedge \forall t, |S_t \cap \mathrm{top}_k(G^m)| \le \frac{k}{k^a} + k^b\right] \ge \mathbb{P}[i \in \hat{V}] \cdot \left(\mathbb{P}\left[\forall t, |S_t \cap \mathrm{top}_k(G^m)| \le \frac{k}{k^a} + k^b\right] - \frac{m}{k^a}\right).$$

Returning to Equation (5):

$$\begin{aligned}&\frac{\frac{k}{k^a}}{\frac{k}{k^a} + k^b} \sum_{i \in \mathrm{top}_k(G^m)} \mathrm{sc}(i, G) \cdot \mathbb{P}\left[i \in \hat{V} \wedge \forall t, |S_t \cap \mathrm{top}_k(G^m)| \le \frac{k}{k^a} + k^b\right] \\ &\ge \frac{\frac{k}{k^a}}{\frac{k}{k^a} + k^b} \sum_{i \in \mathrm{top}_k(G^m)} \mathrm{sc}(i, G) \cdot \mathbb{P}[i \in \hat{V}] \cdot \left(\mathbb{P}\left[\forall t, |S_t \cap \mathrm{top}_k(G^m)| \le \frac{k}{k^a} + k^b\right] - \frac{m}{k^a}\right) \\ &= \frac{\frac{k}{k^a}}{\frac{k}{k^a} + k^b} \left(\mathbb{P}\left[\forall t, |S_t \cap \mathrm{top}_k(G^m)| \le \frac{k}{k^a} + k^b\right] - \frac{m}{k^a}\right) \sum_{i \in \mathrm{top}_k(G^m)} \mathrm{sc}(i, G) \cdot \mathbb{P}[i \in \hat{V}] \\ &= \frac{\frac{k}{k^a}}{\frac{k}{k^a} + k^b} \left(\mathbb{P}\left[\forall t, |S_t \cap \mathrm{top}_k(G^m)| \le \frac{k}{k^a} + k^b\right] - \frac{m}{k^a}\right) \mathbb{E}\left[\sum_{i \in \hat{V}} \mathrm{sc}(i, G)\right].\end{aligned} \tag{7}$$

It remains to show that the ratio of the right hand side of Equation (7) to $\sum_{i \in \mathrm{top}_k(G^m)} \mathrm{sc}(i, G)$ goes to 1. According to Lemma 1,

$$\lim_{k \to \infty} \mathbb{P}\left[\forall t, |S_t \cap \mathrm{top}_k(G^m)| \le \frac{k}{k^a} + k^b\right] = 1.$$

In addition, as $m = o(k^a)$, $\lim_{k \to \infty} \frac{m}{k^a} = 0$. Furthermore, $\frac{k^{1-a}}{k^{1-a} + k^b} = \frac{1}{1 + k^{b-(1-a)}}$, and since $1 - a > b$, $\lim_{k \to \infty} \frac{k^{1-a}}{k^{1-a} + k^b} = 1$. We conclude that

$$\lim_{k \to \infty} \frac{\frac{\frac{k}{k^a}}{\frac{k}{k^a} + k^b} (\mathbb{P}[\forall t, |S_t \cap \mathrm{top}_k(G^m)| \le \frac{k}{k^a} + k^b] - \frac{m}{k^a}) \mathbb{E}\left[\sum_{i \in \hat{V}} \mathrm{sc}(i, G)\right]}{\sum_{i \in \mathrm{top}_k(G^m)} \mathrm{sc}(i, G)} = \frac{\mathbb{E}\left[\sum_{i \in \hat{V}} \mathrm{sc}(i, G)\right]}{\sum_{i \in \mathrm{top}_k(G^m)} \mathrm{sc}(i, G)}.$$

Lemma 4 shows that, for every sampled graph G^m, the ratio on the right hand side approaches 1 as $k \to \infty$ (recall that we have dropped the conditioning on G^m to simplify notation). This proves Equation (4), and thus completes the theorem's proof. $\square$

B Proof of Theorem 2

Lemma 5. *$|P| \leq m$, and thus $\frac{|S|}{k+m} \leq 1$.*

Proof. Recall that for the purposes of computing $\mathrm{top}_k(G^m)$, ties are broken lexicographically. This implies that, for a given $i \notin \mathrm{top}_k(G^m)$, the only way for i to manipulate itself into that set would be to reduce weights on outgoing edges to some of the top k vertices. It can reduce its outgoing weights to at most m vertices; thus, any vertex that makes it into the top k after this manipulation must have been in the top $k+m$ to begin with, where $k+m$ is defined with respect to the tie-breaking order. We conclude that there cannot be more than m vertices that can enter $\mathrm{top}_k(G^m)$ by reducing their outgoing weights. ☐

Lemma 6. *For $i \in S$, $\mathbb{P}[i \in \mathcal{M}^{cs}(G)] = \frac{k}{k+m}$.*

Proof. Some k-subset of S is selected with probability $\frac{|S|}{k+m}$. Given that some k-subset of S is selected, the probability that $i \in S$ is selected is $\frac{k}{|S|}$. Thus,

$$\mathbb{P}[i \in \mathcal{M}^{cs}(G)] = \frac{\mid S \mid}{k+m} \cdot \frac{k}{\mid S \mid} = \frac{k}{k+m}.$$

☐

Lemma 7. *No vertex $i \in N$ can manipulate its way into S.*

Proof. A manipulation placing i into P implies the existence of some manipulation to place i in $\mathrm{top}_k(G^m)$ as well: i was initially in P. A manipulation placing i into $\mathrm{top}_k(G^m)$: i was initially in P. ☐

Proof of Theorem 2. We first show that Credible Subset is impartial. By the definition of the set P and Lemma 7, $\mathbb{P}[i \in S]$ is independent of the outgoing edges of i in the underlying graph. Moreover, by Lemma 6, for any two underlying graphs G and G' that differ only in the weights of the outgoing edges from i,

$$\mathbb{P}[i \in \mathcal{M}^{cs}(G) \mid i \in S] = \frac{k}{k+m} = \mathbb{P}[i \in \mathcal{M}^{cs}(G') \mid i \in S],$$

and

$$\mathbb{P}[i \in \mathcal{M}^{cs}(G) \mid i \notin S] = 0 = \mathbb{P}[i \in \mathcal{M}^{cs}(G') \mid i \notin S].$$

It follows that

$$\mathbb{P}[i \in \mathcal{M}^{cs}(G)] = \mathbb{P}[i \in \mathcal{M}^{cs}(G')].$$

We next show that Credible Subset $\frac{k}{k+m}$-approximates Vanilla. Notice that Credible Subset samples from $\mathcal{G}^m$, just as Vanilla does. In addition, for a fixed sampled graph $G^m \sim \mathcal{G}^m$, Vanilla outputs $\mathrm{top}_k(G^m)$. Thus, for every underlying graph G, the approximation ratio given by Credible Subset is

$$\begin{aligned}\frac{\mathrm{sc}(\mathcal{M}^{cs}, G)}{\mathrm{sc}(\mathcal{M}^{v}, G)} &= \frac{\sum_{G^m} \mathbb{P}[G^m] \cdot \sum_{i \in N} \mathbb{P}[i \in \mathcal{M}^{cs}(G) | \; G^m] \cdot \mathrm{sc}(i, G)}{\sum_{G^m} \mathbb{P}[G^m] \cdot \sum_{i \in N} \mathbb{P}[i \in \mathcal{M}^{v}(G) | \; G^m] \cdot \mathrm{sc}(i, G)} \\ &\geq \frac{\sum_{G^m} \mathbb{P}[G^m] \cdot \sum_{i \in N} \mathbb{I}[i \in \mathrm{top}_k(G^m)] \cdot \frac{k}{k+m} \cdot \mathrm{sc}(i, G)}{\sum_{G^m} \mathbb{P}[G^m] \cdot \sum_{i \in N} \mathbb{I}[i \in \mathrm{top}_k(G^m)] \cdot \mathrm{sc}(i, G)} = \frac{k}{k+m},\end{aligned}$$

where the second transition follows from Lemma 6, and $\mathbb{I}[E]$ is an indicator variable that takes that value 1 if the event E is true and 0 if E is false. $\square$

C Proof of Theorem 3

We require the following straightforward probabilistic lemma.

Lemma 8. *Let $c \in (0, 1/4)$. Suppose n^c distinct elements are drawn from a universe of size n. Suppose this experiment is repeated n^c times, and let the selected set in round t be denoted N_t. Then, with high probability, $N_t \cap N_{t'} = \emptyset$, for all $t \neq t'$.*

Proof. Let $X = \{1, \ldots, n^c\}$. Using the union bound,

$$\mathbb{P}[\exists i \neq j \in X,\ N_i \cap N_j \neq \emptyset] \leq \binom{n^c}{2} \cdot \mathbb{P}[N_1 \cap N_2 \neq \emptyset].$$

Suppose that we have drawn the items in N_1, and now we draw the n^c items in N_2 one by one. The probability that an item in N_2 coincides with an item of N_1 is at most $\frac{n^c}{n-n^c}$, where the denominator accounts for the fact that some of the items in N_2 may have already been chosen. Using the union bound again,

$$\mathbb{P}[N_1 \cap N_2 \neq \emptyset] \leq \frac{n^{2c}}{n - n^c},$$

and overall

$$\mathbb{P}[\exists i \neq j \in X,\ N_i \cap N_j \neq \emptyset] \leq \binom{n^c}{2} \frac{n^{2c}}{n - n^c} \leq \frac{2n^{4c}}{n} \xrightarrow{n \to \infty} 0.$$

$\square$

Proof of Theorem 3. Let $\mathcal{M}$ be an impartial mechanism. Consider a set $X \subset N$ of size m. We will build up a matching between X and $N \setminus X$, such that the probability $\mathcal{M}$ samples the edge $(\mu(i), i)$ is small (roughly m/n) for all i. This will imply that $\mathcal{M}$ will have to select i with similar probability on two graphs which differ only in the weight of the edge $(\mu(i), i)$.

Select an arbitrary vertex and label it 1. Let $\mu(1) = \mathrm{argmin}_j \mathbb{P}[\mathcal{M} \text{ samples } (j, 1)]$ (the vertex with the smallest probability of $(j, 1)$ being sampled by $\mathcal{M}$). Let $q_1 = \mathbb{P}[\mathcal{M} \text{ samples } (\mu(1), 1)]$; note that $q_1 \leq \frac{m}{n-1}$ by a simple averaging argument. Then, for each $i \in [2, \ldots, m]$, select another arbitrary vertex and label it i such that $i \notin \{1, \ldots, i-1\} \cup \{\mu(1), \ldots, \mu(i-1)\}$, and let

$$\mu(i) = \mathrm{argmin}_{j \notin \{1, \ldots, i\} \cup \{\mu(1), \ldots, \mu(i-1)\}} \mathbb{P}[\mathcal{M} \text{ samples } (j, i)],$$

be the vertex such that $(\mu(i), i)$ has the smallest probability of being sampled by $\mathcal{M}$ which is not already part of the matching, and

$$q_i = \mathbb{P}[\mathcal{M} \text{ samples } (\mu(i), i)]$$

be that probability. Note that $q_i \leq \frac{m}{n-2(i-1)-1}$, else the expected number of edges incident to i would be larger than m.

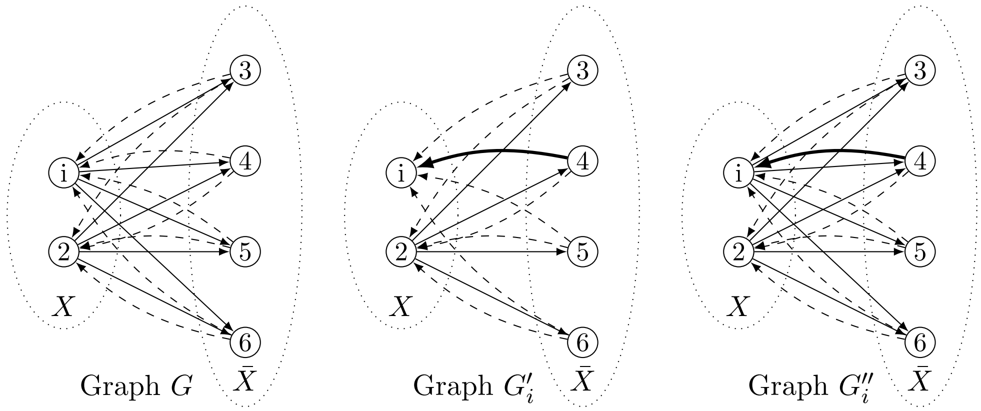

Figure 1: Example of the graphs G, G'_i, G''_i where $i = 1$, $\mu(i) = 4$, $X = \{i, 2\}$. Solid lines represent edges of weight 1, dashed lines edges of weight ϵ, and thick lines edges of weight M. Edges not present have weight 0. G and G''_i differ only on the weight of edge $(4, i)$; G''_i and G'_i differ only on the weight of outgoing edges from i.

Now, we construct an underlying graph G be defined using the following weights:

$$w_G(i,j) = \begin{cases} 1 & i \in X, j \notin X \\ \epsilon \ll \frac{1}{m} & i \notin X, j \in X \\ 0 & \text{otherwise} \end{cases}$$

For each $i \in X$, let the graph G'_i on n vertices be as follows:

$$w_{G'_i}(j,j') = \begin{cases} M \gg 1 & j = \mu(i), j' = i \\ 1 & j \in X, j \neq i, j' \notin X \\ \epsilon \ll \frac{1}{m} & j \notin X, j' \in X, (j,j') \neq (\mu(i), i) \\ 0 & \text{otherwise} \end{cases}$$

Notice that G'_i differs from G in two ways: it has one high-weight edge to i, and the outgoing edges from i have weight 0 rather than weight 1. For an illustration, see Figure 1.

We begin by showing that

$$\mathrm{sc}(\mathcal{M}^v, G) \geq |X|k(1 - o(1)). \tag{8}$$

To prove (8), denote the set of vertices adjacent to a set Y in the sampled graph G^m by $\mathcal{N}_{G^m}(Y)$. Notice that the vertices $j \in \mathcal{N}_{G^m}(X)$ have strictly higher sampled ratings than all other vertices in G^m. Moreover, $|\mathcal{N}_{G^m}(X)| \leq k$, so VANILLA will select all $j \in \mathcal{N}_{G^m}(X)$. Thus,

$$\begin{aligned} \mathrm{sc}(\mathcal{M}^v, G) &= \sum_j \mathbb{P}[j \in \mathrm{top}_k(G^m)]\mathrm{sc}(j,G) \geq \sum_{j \notin X} \mathbb{P}[j \in \mathrm{top}_k(G^m)]\mathrm{sc}(j,G) \\ &\geq \sum_{j \notin X} \mathbb{P}[j \in \mathcal{N}_{G^m}(X)]\mathrm{sc}(j,G) \geq |X| \sum_{j \notin X} \mathbb{P}[j \in \mathcal{N}_{G^m}(X)] = |X| \cdot \mathbb{E}\left[|\mathcal{N}_{G^m}(X)|\right] \\ &\geq |X|(k(1 - o(1))), \end{aligned}$$

where the penultimate transition follows from Lemma 8 and the assumption that $c \in (0, \frac{1}{4})$ and $m \leq n^c$.

Next, we claim that

$$\mathrm{sc}(\mathcal{M}^v, G'_i) \geq M. \tag{9}$$

Let G^m denote the sampled graph. Then, notice that there is a trivial upper bound on the size of $|\mathcal{N}_{G^m}(X \setminus \{i\})|$:

$$|\mathcal{N}_{G^m}(X \setminus \{i\})| \leq m(|X| - 1) = k - m. \tag{10}$$

Therefore,

$$\begin{aligned}
\mathrm{sc}(\mathcal{M}^v, G'_i) &= \sum_j \mathbb{P}[j \in \mathrm{top}_k(G^m)]\mathrm{sc}(j, G'_i) \geq M \cdot \mathbb{P}[i \in \mathrm{top}_k(G^m)] \\
&\geq M \cdot \mathbb{P}[X \subset \mathrm{top}_k(G^m)] = M \cdot \mathbb{P}[|\mathcal{N}_{G^m}(X \setminus \{i\})| \leq k - m] = M.
\end{aligned}$$

The fourth transition follows from the observation that the only vertices with nonzero sampled ratings are in $X \cup \mathcal{N}_{G^m}(X \setminus \{i\})$ (which implies Vanilla will select all of them, if there aren't more than k), and the final equality comes from from (10).

Now, we revisit the impartial mechanism $\mathcal{M}$. We show the probability i is selected by $\mathcal{M}$ in G cannot be too different from the probability i is selected by $\mathcal{M}$ in G'_i. Let $p_i = \mathbb{P}[i \in \mathcal{M}(G)]$. Consider the "intermediate" graph G''_i such that

$$w_{G''_i}(j, j') = \begin{cases} M \gg 1 & j = \mu(i), j' = i \\ 1 & j \in X, j' \notin X \\ \epsilon \ll \frac{1}{m} & j \notin X, j' \in X, (j, j') \neq (\mu(i), i) \\ 0 & \text{otherwise} \end{cases}$$

That is, G''_i is the graph G with the added heavy-weight edge to i, or the graph G'_i with the outgoing edges from i set to 1. See Figure 1 for an illustration.

Let G^m be the graph sampled by $\mathcal{M}$. If $(\mu(i), i) \notin E(G^m)$, $\mathcal{M}$ cannot distinguish between G and G''_i, and thus must select i with the same probability in those cases. Then, by impartiality, $\mathcal{M}$ must select i with equal (unconditional) probability in G'_i, G''_i, since they differ only in the outgoing edges from i.

In more detail, let us denote $p_i = \mathbb{P}[i \in \mathcal{M}(G)]$. We have

$$\begin{aligned}
p_i &= \mathbb{P}[i \in \mathcal{M}(G) \mid (\mu(i), i) \in E(G^m)]\mathbb{P}[(\mu(i), i) \in E(G^m)] \\
&\quad + \mathbb{P}[i \in \mathcal{M}(G) \mid (\mu(i), i) \notin E(G^m)]\mathbb{P}[(\mu(i), i) \notin E(G^m)] \\
&= \mathbb{P}[i \in \mathcal{M}(G) \mid (\mu(i), i) \in E(G^m)]\mathbb{P}[(\mu(i), i) \in E(G^m)] \\
&\quad + \mathbb{P}[i \in \mathcal{M}(G) \mid (\mu(i), i) \notin E(G^m)](1 - \mathbb{P}[(\mu(i), i) \in E(G^m)]).
\end{aligned}$$

Then, we explicitly write p_i in terms of q_i:

$$p_i = \mathbb{P}[i \in \mathcal{M}(G) \mid (\mu(i), i) \in E(G^m)]q_i + \mathbb{P}[i \in \mathcal{M}(G) \mid (\mu(i), i) \notin E(G^m)]\,(1 - q_i)\,.$$

Therefore,

$$\begin{aligned}
\mathbb{P}[i \in \mathcal{M}(G''_i) \mid (\mu(i), i) \notin E(G^m)] &= \mathbb{P}[i \in \mathcal{M}(G) \mid (\mu(i), i) \notin E(G^m)] \\
&= \frac{p_i - q_i \mathbb{P}[i \in \mathcal{M}(G) \mid (\mu(i), i) \in E(G^m)]}{(1 - q_i)} \leq \frac{p_i}{(1 - q_i)}.
\end{aligned}$$

We can use this inequality to derive an upper bound on the probability that $i \in \mathcal{M}(G_i'')$:

$$\begin{aligned}\mathbb{P}[i \in \mathcal{M}(G_i'')] &= (1-q_i)\mathbb{P}[i \in \mathcal{M}(G_i'')|(\mu(i),i) \notin E(G^m)] + q_i\mathbb{P}[i \in \mathcal{M}(G_i'')|(\mu(i),i) \in E(G^m)] \\ &\leq (1-q_i)\frac{p_i}{1-q_i} + q_i = p_i + q_i.\end{aligned}$$

Then, by impartiality, $\mathbb{P}[i \in \mathcal{M}(G_i')] = \mathbb{P}[i \in \mathcal{M}(G_i'')] \leq p_i + q_i$. It follows that

$$\begin{aligned}\frac{\mathrm{sc}(\mathcal{M}, G_i')}{\mathrm{sc}(\mathcal{M}^v, G_i')} &\leq \frac{(p_i+q_i)(M+(k-1)(|X|-1)) + (1-p_i-q_i)k(|X|-1)}{M} \\ &= p_i + q_i + \frac{(p_i+q_i)(k-1)(|X|-1) + (1-p_i-q_i)k(|X|-1)}{M} \\ &\leq p_i + q_i + \frac{(p_i+q_i)k(|X|-1) + (1-p_i-q_i)k(|X|-1)}{M} = p_i + q_i + \frac{k(|X|-1)}{M}\end{aligned} \tag{11}$$

where the first inequality comes from a simple calculation of scores, Equation (9), and the bound $p_i + q_i \geq \mathbb{P}[i \in \mathcal{M}(G_i')]$.

On the other hand, let $p = \frac{\sum_{i \in X} p_i}{m}$. Then

$$\begin{aligned}\frac{\mathrm{sc}(\mathcal{M}, G)}{\mathrm{sc}(\mathcal{M}^v, G)} &\leq \frac{(k - \sum_{i\in X} p_i)|X| + \epsilon(n-|X|)\sum_{i\in X} p_i}{(1-o(1))|X|k} \\ &= \frac{(k - \sum_{i\in X} p_i)m + \epsilon(n-m)\sum_{i\in X} p_i}{(1-o(1))mk} = \frac{(k-pm)m + \epsilon\,(n-m)\,pm}{(1-o(1))mk} \\ &= \frac{\left(1-\frac{pm}{k}\right) + \epsilon\,(n-m)\,\frac{p}{k}}{(1-o(1))} \leq \frac{\left(1-\frac{pm}{k}\right) + \epsilon n\frac{p}{k}}{(1-o(1))}.\end{aligned} \tag{12}$$

Now, some $p_i \leq p$, by a simple averaging argument; consider that i. In the construction of μ above, we showed the upper bound $q_i \leq \frac{m}{n-2(i-1)-1}$ on the probability that $(\mu(i), i)$ is sampled by $\mathcal{M}$. Notice that the approximation ratio for $\mathcal{M}$ is at most

$$\begin{aligned}\alpha &\leq \min\left\{p_i + q_i + \frac{k(|X|-1)}{M}, \frac{\left(1-\frac{pm}{k}\right)+\epsilon n\frac{p}{k}}{(1-o(1))}\right\} \\ &\leq \min\left\{p + q_i + \frac{k(|X|-1)}{M}, \frac{\left(1-\frac{pm}{k}\right)+\epsilon n\frac{p}{k}}{(1-o(1))}\right\},\end{aligned}$$

by (11) and (12). Since ϵ is arbitrarily small, M is arbitrarily large, and $q_i = o(1)$,

$$\alpha \leq \min\left\{p, \left(1 - \frac{pm}{k}\right)\right\} + o(1).$$

We derive an upper bound on the minimum by equalizing the two expressions and solving for p, which yields $p = \frac{k}{k+m}$. It follows that $\alpha \leq \frac{k}{k+m} + o(1)$. □