ECAI 2024
U. Endriss et al. (Eds.)

doi:10.3233/FAIA240882

Online Friends Partitioning Under Uncertainty

Saar Cohen[a] and Noa Agmon[a]

[a]Department of Computer Science, Bar-Ilan University, Israel; saar30@gmail.com, agmon@cs.biu.ac.il

Abstract. We study the *friendship-based online coalition formation* problem, in which agents that appear one at a time should be partitioned into coalitions, and an agent's utility for a coalition is the number of her neighbors (i.e., *friends*) within the coalition. Unlike prior work, agents' friendships may be *uncertain*. We analyze the desirability of the resulting partition in the common term of *optimality*, aiming to maximize the social welfare. We design an online algorithm termed *Maximum Predicted Coalitional Friends* (MPCF), which is enhanced with predictions of each agent's number of friends within any possible coalition. For common classes of random graphs, we prove that MPCF is optimal, and, for certain graphs, provides the *same* guarantee as the best known competitive algorithm for settings *without* uncertainty.

1 Introduction

For the social dinner at ECAI 2024, the organizers have reserved the best banquet venue of the city. When you arrive, you see multiple tables of various capacities that have been set to accommodate the participants, yet some seats are already taken. You would like to share your table with some friends you have not seen in a long time. However, you face uncertainty about your friendships due to shifts in collaborations and the virtual nature of recent years. Given this uncertainty, the organizers would like to ensure that each attendee is assigned to a table with as many of his friends as possible, so he would have a good experience during the event. A similar scenario was considered by Bilò et al. [17]. Additional real-life examples include academic research [2] and international unions [39], where *agents* perform activities in *coalitions* rather than on their own. Such scenarios fall within the phenomenon of coalition formation, which is noticed in our social, economic, and political life.

In this paper, we introduce and explore a model capturing such real-life situations. A popular framework for studying coalition formation is that of *hedonic games* [32], which disregards externalities, i.e., agents' utilities depend solely on the coalition they are part of. The outcome of such games is a set of disjoint coalitions (hereafter, *partition*). In our model, agents' friendships may be represented by an unweighted and undirected graph, where an agent's utility for a coalition is the number of her neighbors (i.e., *friends*) within the coalition, i.e., the agent's *degree* in the graph induced by the coalition. As in our example, there are many contexts where it is more realistic to assume that coalitions may have a limited *physical* space and agents arrive over time. Thus, we assume an upper bound on the size of each coalition. Then, a central authority (i.e., an *online* algorithm) has to *immediately* and *irrevocably* decide whether to add an arriving agent into an existing coalition or to create a new one containing, at this moment, only her.

Unlike prior works where arriving agents are assumed to reveal their exact friendships with previously disclosed agents [24, 25, 37], we consider cases where friendships are *uncertain*, and design an online algorithm augmented with a *coalitional friends predictor*, i.e., an oracle that predicts an agent's number of friends within a given coalition. Not only that such a predictor is simple and easy to interpret, it is also useful: Vertices' degree information has been previously employed in heuristic and approximation algorithms for other graph problems (e.g., maximum independent set [38]). Since an agent's degree is simply her frequency in the union of all edges, a predictor can also be readily attained via estimation methods of elements' frequencies in a dataset, as used for certain data analysis problems [40].

The quality of a partition has been measured by various solution concepts, such as stability and optimality [21]. In this paper, we explore the objective of *maximizing the (utilitarian) social welfare*, defined as the sum of the agents' utilities from a given partition. We study a simple algorithm called *Maximum Predicted Coalitional Friends* (MPCF), which assigns agents to coalitions greedily with respect to the predictor. For general graphs, MPCF returns a partition with a high social welfare even when the predictions are slightly off. For illustrating MPCF's good performance, we follow a vast trend of research on average-case analysis (see, e.g., [41, 49]), and analyze MPCF also under a very natural and common random graph model: the Chung-Lu-Vu (CLV) model [26], which generalizes the well-known Erdős-Rényi model [35]. When the expected degree of each agent in the subgraph induced by any coalition is used as a prediction, we prove that MPCF stochastically dominates *any* other algorithm for graphs drawn from the CLV model and analyze its expected social welfare. For *deterministic* graphs, we also show that MPCF has the *same* competitive ratio as the best known algorithm introduced by Flammini *et al.* [37]. Surprisingly, our work illustrates that uncertainty about the agents' preferences gives rise to an optimal algorithm for a very natural random graph model, unlike scenarios *without* uncertainty where the best known algorithm by Flammini et al. [37] is only *almost* optimal.

2 Related Work

Our research can be viewed as *additively separable hedonic games* (ASHGs) [18], where preferences are binary and symmetric under the restriction that each coalition's size is bounded. In ASHGs, agents' preferences are encoded using weighted graphs, while coalition evaluation is based on the summation of members' valuations within the coalition. Hedonic games, introduced by Dreze and Greenberg [32], have since been extended to various solution concepts like stability and fairness [9, 55]. Designing computationally feasible classes of hedonic games is a major concern, resulting in different representations. Cardinal hedonic games like ASHGs, based on weighted graphs [10, 18], offer space efficiency compared to ordinal representations [19, 33]. ASHGs encompass friend-oriented hedonic games [30], where friends and enemies are distinguished through two possible weights. We focus on a subclass of these games, where an agent can express for any other agent if she is a neutral or a friend.

Partitions in hedonic games are typically measured in terms of stability and optimality. While in [12, 18] properties guaranteeing the existence of stable partitions in ASHGs were supplied, their computational aspects were studied in [6, 11]. Unlike most prior studies on friend-oriented hedonic games that focus on stability notions [16, 22, 42, 45, 47], we consider measures of optimality. Specifically, in our work we concentrate on maximizing social welfare, as also studied in *offline* settings [8, 23, 34, 51]. In fact, Levinger et al. [46] explain why our problem is already computationally hard in *offline* setting, and they thus provide poly-time approximation algorithms. Bilò et al. [17] show that social welfare is hard to approximate even in a restricted *offline* variant of our problem with fixed-size coalitions, which they complement with poly-time approximation algorithms. In *online* settings, Flammini et al. [37] study social welfare maximization in cases similar to standard *symmetric* ASHGs, where agents arrive along with their incident edges. Bullinger and René [24] also explored a similar setup, and recently they examined whether various stability notions can be attained in such online settings [25]. Cohen and Agmon [28] study a coalitional variant of online task allocation, where each coalition is given a task and evaluated on its members' skills, yielding a multi-dimensional utility structure, unlike our work.

In contrast, we aim at maximizing social welfare in *online* hedonic games when agents' friendships (and thus their utilities) are *uncertain* (i.e., edges are *not* revealed). Cohen and Agmon [27] were the first to consider agents' uncertainty about friendships, yet in *offline* settings under the restricted Erdős-Rényi model [35]. However, we consider the most general setting of uncertain friendships, while exhibiting our algorithm's good performance under the Chung-Lu-Vu (CLV) model [26] that generalizes the Erdős-Rényi model. Our results also provide a *major* contribution compared to prior work on online coalition formation. For *non-negative* utilities, the greedy algorithm by Flammini et al. [37] is only *almost* optimal when agents reveal upon arrival their valuations to previously arrived agents. However, in our *new* setting where valuations are *not* revealed, our algorithm is *optimal* for the very natural CLV random graph model.

Our work is also closely related to a recently popular trend of augmenting online algorithms with predictions, with the aim of bypassing the worst-case lower bounds of online problems caused by the uncertainty of the future. Though the idea of using advice to obtain semi-online algorithms is not new [20], Munoz and Vassilvitskii [52] propose to use a predictor oracle for improving revenue optimization in auctions by setting a good reserve (or minimum) price, and Lykouris and Vassilvitskii [48] consider the online caching problem with predictions. These works led to a series of learning augmented results in various fields (e.g., clustering [31], ski-rental [3]). To the best of our knowledge, we are the first to introduce predictions to online coalition formation problems, including a scheme that is stochastically optimal for realistic and natural random graphs.

Predictions about agents' utilities have been proven to yield significantly improved approximations in other problems. For instance, Banerjee et al. [13, 14] augment online fair division problems with predictions about each agent's value for *all* items. Similarly to a recent work on online bipartite matching by Aamand et al. [1], we also use predictions derived from agents' expected degrees. However, in the above works agents or items arrive along with their incident edges, while in ours edges they are not revealed. Further, while they use predictions about an agent's utility from *the full graph*, the oracle used by our MPCF algorithm predicts an agent's utility from *the subgraph induced by a given coalition*. Moreover, the algorithm by Aamand et al. [1] greedily matches an arriving agent with a *minimum* predicted degree neighbor that is yet to be covered. In contrast, MPCF assigns an arriving agent to a coalition with *maximum* predicted number of friends that has the *minimum* total expected degree among such coalitions. In other graph problems, other kinds of predictions have also been used (See, e.g., [4, 43]).

3 Online Partitioning of Friends

We consider the problem of partitioning a finite set $N = \{1, \ldots, n\}$ of n agents within an undirected social network $G = (N, E)$ into coalitions, where an agent benefits from the arrival of another agent into her coalition so long as they are neighbors (i.e., friends). Agent i's preferences can thus be succinctly represented by a *(binary) cardinal utility function* $v_i : N \to \{0, 1\}$ with $v_i(i) = 0$ that satisfies $v_i(j) = 1$ if agent $j \neq i$ is a friend of agent i (i.e., $(i, j) \in E$); otherwise, $v_i(j) = 0$. We denote by $\mathcal{N}_i$ the set of coalitions agent i belongs to, i.e., $\mathcal{N}_i = \{C \subseteq N : i \in C\}$. Agent i's utility can be *additively* aggregated to preferences over each coalition $C \in \mathcal{N}_i$ via $v_i(C) = \sum_{j \in C} v_i(j)$. Such representation allows us to compare agents' utilities such that a certain cardinal value expresses the same intensity of a preference for all agents. In fact, our model corresponds to *additively separable hedonic games* (ASHGs) with symmetric and binary preferences [23].

An outcome is thus a *partition* π of N into disjoint coalitions, where $|\pi|$ denotes the number of its coalitions. Let $\pi(i)$ be the coalition $C \in \pi$ such that $i \in C$. Hence, for a partition π, $\pi(i) \succeq_i \pi'(i)$ iff $v_i(\pi) \geq v_i(\pi')$, where $v_i(\pi) = v_i(\pi(i))$ is the utility i receives from a partition π. We focus on real-life scenarios where the size of each coalition is bounded. For instance, regarding our banquet example, each table can accommodate a limited number of employees. Hence, for a positive integer α, we consider partitions π that are α-*bounded*, i.e., $|C| \leq \alpha$ for every coalition $C \in \pi$. We assume that $\alpha \geq 2$ as the case $\alpha = 1$ is trivial. We denote by Π_α the collection of all α-bounded outcomes. For an integer $n > 0$, we henceforth denote $[n] := \{1, \ldots, n\}$ where $[0] = \{0\}$.

We evaluate the quality of partitions by measures of optimality. We regard partitions that *maximize the (utilitarian) social welfare*, which is defined as the sum of all agents' utilities for that partition. Formally, for a coalition $C \subseteq \mathcal{N}$, let $\mathcal{SW}(C) = \sum_{i \in C} v_i(C)$ be the *social welfare of* C. The *social welfare of a partition* π is then $\mathcal{SW}(\pi) = \sum_{C \in \pi} \mathcal{SW}(C) = \sum_{i \in \mathcal{N}} v_i(\pi)$. Hence, a partition π is *welfare-optimal* if it maximizes the social welfare amongst all possible partitions, i.e., $\pi \in \arg\max_{\pi' \in \Pi_\alpha} \mathcal{SW}(\pi')$.

In our *online* model, agents appear one at a time in the order $1, \ldots, n$. At each time t, an online algorithm $\mathcal{A}$ shall produce a partial partition π^t of the agents *who arrived until time* t, without any knowledge regarding future agents. Upon the arrival of agent t, $\mathcal{A}$ should *immediately* and *irrevocably* decide whether to insert her to an existing coalition in π^{t-1} or create a new coalition $\{t\}$. For an integer $\alpha \geq 2$, the central goal of an online algorithm $\mathcal{A}$ is thus computing a welfare-optimal α-bounded partition π^n. As the number of agents is not known upfront, from the perspective of an online algorithm, it may be possibly infinite. Prior research considered scenarios where friendships are revealed upon arrival [24, 25, 37]. See [29, Appendix A] for a sample instance in such settings.

Unlike prior studies, and unless stated otherwise, we consider that agents do *not* know their edges to previously arrived agents. Thus,

Algorithm 1 **Maximum Predicted Coalitional Friends**

1: Initialize an empty partition $\pi \leftarrow \emptyset$.
2: **while** an online agent $t \in N$ arrives **do**
3: Set $\pi' = \{C \in \pi : |C| < \alpha \wedge \varphi_t(C) > 0\}$.
4: **if** $|\pi'| > 0$ **then**
5: Set $\mathcal{S} = \arg\max_{C \in \pi'} \varphi_t(C)$ and add agent t to a coalition $C \in \arg\min_{C' \in \mathcal{S}} \sum_{i \in C'} \varphi_i(N)$ (ties broken randomly).
6: **else** Create a new coalition $\{t\}$ and add it to π.
7: **Output:** The partition π.

agents do *not* reveal them upon arrival. In many practical scenarios, it is more realistic to assume that agents have some knowledge about their social connections rather than being completely ignorant. In our banquet example, many companies use social network analysis software to visualize and analyze relationships among employees, which can identify clusters of employees who frequently interact, indicating potential friendships (though those friendships are not certain). For such cases, we thus consider online algorithms augmented with a "coalitional friends predictor" $\varphi_i : 2^N \to \mathbb{R}_{\geq 0}$ for each agent $i \geq 2$, which are possibly stochastic, inferred from additional knowledge about the graph or machine-learned from past data. We use such notation to stress that the predictors can operate on any coalition, but in practice they only operate on subsets of previously disclosed agents. Intuitively, the predictor captures agent i's *uncertainty* on her friendships: it is an oracle that, given any coalition $C \subseteq N$, predicts the number of agent i's friends within C. In particular, $\varphi_i(N)$ predicts agent i's degree in the full graph. For the predictor to be well-defined, we assume that $\varphi_i(C) = \sum_{j \in C} \varphi_i(\{j\})$ for any coalition $C \subseteq N$.

4 Maximizing Social Welfare

In this section, we present our online algorithm called *Maximum Predicted Coalitional Friends* (**MPCF**), which assigns agents to coalitions greedily with respect to the predictor. Our MPCF algorithm (Algorithm 1) uses the predictors to greedily assign an arriving agent t to a coalition $C \in \pi^{t-1}$ of size less than α that contains the maximum predicted number of friends, yielding the maximum *positive* increase in the current partition's welfare. If multiple such coalitions exist, the algorithm assigns agent t to the coalition C whose total predicted expected degree within the entire graph is *minimal*, i.e., C minimizes $\sum_{i \in C} \varphi_i(N)$. Intuitively, coalitions with *low* total degrees should be filled as early as possible as we will have more chances to assign agents to coalitions with *higher* total degrees. If no such coalition exists, then MPCF creates a new singleton coalition $\{t\}$.

As an algorithm, MPCF is simple, but our main novelty is the analysis of its behaviour with respect to the quality and choice of the predictors. For general graphs, we show in Section 4.1 that MPCF maintains a high social welfare even when the predictors are *noisy*. For depicting its good performance, we prove in Section 4.2 that MPCF is *optimal* for a very natural random graph model and analyze its expected social welfare. In the sequel, we denote the coalition to which agent i is assigned by an online algorithm $\mathcal{A}$ as $\mathcal{A}_i(G)$. For brevity, we hereafter denote MPCF by $\mathcal{A}^\star$.

4.1 Robustness of MPCF to Noisy Predictors

Generally, we would like the algorithm to perform well when the predictions are decent (or even accurate), yet maintain reasonable performance even when the predictions are slightly *noisy*. For *general* graphs, we thus show that MPCF will return a partition with a high social welfare even when the predictions are slightly off. Formally, consider two coalitional friends predictors $\varphi = (\varphi_i)_{i \in N}$ and $\varphi' = (\varphi'_i)_{i \in N}$. Given any pair of agents $j \neq i$, note that $\varphi_i(\{j\})$ and $\varphi'_i(\{j\})$ predict whether agents i and j are friends or not, where a higher prediction may indicate a stronger potential for friendship. However, for each agent i, φ_i and φ'_i may disagree upon their predictions of agent i's friendship with any other agent j (i.e., $\varphi_i(\{j\}) \neq \varphi'_i(\{j\})$ may hold) and thus induce different orderings on the agents who arrived *before* agent i. Note that the ordering is done by sorting the agents who arrived before agent i based on the predictions of friendship with agent i in descending order. Hence, let $\Delta(\varphi_i, \varphi'_i)$ be the *minimal* set of agents that should be removed such that the two predictors φ_i and φ'_i will induce the same ordering over the remaining agents in $[i-1]$.

Let $\Delta(\varphi, \varphi') := \cup_{i \in N} \Delta(\varphi_i, \varphi'_i)$. Next, we give an upper bound on the difference between the social welfares incurred by executing MPCF with the predictors φ_i and φ'_i. Our derivation of the upper bound leverages the fact that the removal of a single agent in the graph cannot yield a better social welfare, as proven in Lemma 1.

Lemma 1. *Let $G = (N, E)$ be a graph and consider an agent $j \in N$. For each agent i, consider a coalitional friends predictor φ_i. Then, MPCF with predictors φ_i when executed on $G' = (N \setminus \{j\}, E)$ will yield a partition whose social welfare is at most the one incurred when executed on G.*

Proof. Let $N_G(i)$ be the neighborhood of any agent $i \in N$ in the graph G, only including agents that are *not* in a coalition of size α upon the arrival of agent i. We will prove by induction that $N_{G'}(i) \subseteq N_G(i)$ for any agent i. In base case where the first agent arrives (i.e., agent 1), if $j \notin N_G(1)$, then $N_{G'}(1) = N_G(1)$; otherwise, $N_{G'}(i) = N_G(i) \setminus \{j\}$.

For the inductive step, assume that $i > 1$ and $N_{G'}(h) \subseteq N_G(h)$ for any agent $h < i$. We assume by contradiction that there exists an agent h' such that $h' \in N_{G'}(i)$, but $h' \notin N_G(i)$. Hence, prior to agent i's arrival, h' was assigned by MPCF when executed on G to a coalition that is of size α at time i, but not when MPCF was executed on G'. Let C be the coalition to which agent h' was assigned when MPCF was executed on G. Since $h' \in N_{G'}(i)$, it must be that $h' \in N_G(f)$ for any $f \in C$ since when MPCF was executed on G' it has not yet assigned h' at time i and $h' \in N_G(f)$ for any $f \in C$. Let C'_f be the coalition to which each agent $f \in C$ was assigned when MPCF was executed on G'. Since MPCF had to decide whether to assign agent f to coalition C or coalition C'_f when executed on G', it must be that either $\varphi_f(C'_f) > \varphi_f(C)$ or $\varphi_f(C'_f) = \varphi_f(C)$ and $\sum_{f' \in C'_f} \varphi_{f'}(N) < \sum_{f'' \in C} \varphi_{f''}(N)$. Further, by the inductive hypothesis, for each $f \in C$, each $f' \in C'_f$ and $f'' \in C$ with $f', f'' \in N_{G'}(f)$ also satisfy $f', f'' \in N_G(f)$. Thus, since MPCF assigned agent f to coalition C instead of coalition C'_f, we infer that either $\varphi_f(C'_f) < \varphi_f(C)$ or $\varphi_f(C'_f) = \varphi_f(C)$ and $\sum_{f' \in C'_f} \varphi_{f'}(N) > \sum_{f'' \in C} \varphi_{f''}(N)$. This constitutes a contradiction.

Note that each agent's utility from a given partition can be at most her number of friends in the graph. Thus, since $|N_{G'}(i)| \leq |N_G(i)|$ by our proof above, we infer the desired statement in Lemma 1. □

In Theorem 2, we provide our upper bound:

Theorem 2. *Let $\mathcal{A}^\star_\varphi(G)$ be the partition generated by MPCF with a predictor φ on a graph G. Then, $\mathcal{SW}(\mathcal{A}^\star_\varphi(G)) - \mathcal{SW}(\mathcal{A}^\star_{\varphi'}(G)) \leq |\Delta(\varphi, \varphi')|(2\alpha - 3)$ for any pair of predictors φ and φ'.*

Proof. Let $\mathcal{A}^\star_\varphi(G)$ be the partition generated by MPCF with a predictor φ on a graph G. Let $G_{-\Delta(\varphi,\varphi')} = (N_{-\Delta(\varphi,\varphi')}, E_{-\Delta(\varphi,\varphi')})$

be the graph resulting from the removal of the agents in $\Delta(\varphi, \varphi')$, where $N_{-\Delta(\varphi,\varphi')} = N \setminus \Delta(\varphi_i, \varphi'_i)$ and $E_{-\Delta(\varphi,\varphi')} = E \cap (N_{-\Delta(\varphi,\varphi')} \times N_{-\Delta(\varphi,\varphi')})$. By invoking Lemma 1 inductively:

$$\mathcal{SW}(\mathcal{A}^\star_{\varphi'}(G)) \geq \mathcal{SW}(\mathcal{A}^\star_{\varphi'}(G_{-\Delta(\varphi,\varphi')})) \tag{1}$$

Next, let $\deg_G(i)$ be agent i's actual degree in the graph G. Consider an agent $j \in \Delta(\varphi, \varphi')$. Note that agent j can receive at most a utility of $\min(\deg_G(j), \alpha - 1)$. Thus, upon the removal of agent j, the utility corresponding to each of agent j's neighbors reduces by at most 1, and thus the partition's social welfare can decrease by at most $\min(\deg_G(j), \alpha - 1) + (\min(\deg_G(j), \alpha - 1) - 1) = 2\min(\deg_G(j), \alpha - 1) - 1$. By invoking those arguments recursively for each agent $j \in \Delta(\varphi, \varphi')$, we have:

$$\begin{aligned} &\mathcal{SW}(\mathcal{A}^\star_{\varphi}(G)) \leq \mathcal{SW}(\mathcal{A}^\star_{\varphi}(G_{-\Delta(\varphi,\varphi')})) + \\ &+ \sum_{j \in \Delta(\varphi,\varphi')} [2\min(\deg_G(j), \alpha - 1) - 1] \leq \\ &\leq \mathcal{SW}(\mathcal{A}^\star_{\varphi}(G_{-\Delta(\varphi,\varphi')})) + |\Delta(\varphi, \varphi')|(2\alpha - 3) \end{aligned} \tag{2}$$

Note that $\mathcal{SW}(\mathcal{A}^\star_{\varphi}(G_{-\Delta(\varphi,\varphi')})) = \mathcal{SW}(\mathcal{A}^\star_{\varphi'}(G_{-\Delta(\varphi,\varphi')}))$ since removing the agents in $\Delta(\varphi, \varphi')$ yields that the predictors φ and φ' induce the same set of predicted friends for each agent, and thus $\mathcal{SW}(\mathcal{A}^\star_{\varphi}(G)) \leq \mathcal{SW}(\mathcal{A}^\star_{\varphi'}(G_{-\Delta(\varphi,\varphi')})) + |\Delta(\varphi, \varphi')|(2\alpha - 3)$ by (2). Combined with (1), we conclude the desired upper bound. □

Given any other predictor φ', we infer that MPCF will still incur a high social welfare for a small enough value of $|\Delta(\varphi, \varphi')|$, which is upper bounded by the number of mispredicted friends of each agent. In the next section, we show that when MPCF is *optimal* under the CLV model when MPCF uses predictors that return the expected number of each agent's friends within a given coalition. This implies that, so long as φ's number of mispredicted friends for each agent i is *small*, MPCF remains still *near-optimal*.

4.2 *Optimality of MPCF for the CLV Model*

For depicting MPCF's good performance and effectively analyzing it, we consider the natural Chung-Lu-Vu (CLV) random graph model [26] that generates graphs with arbitrary expected degree sequences, generalizing the well-studied known i.i.d. model (see, e.g., [36]). Our analysis stresses the applicability of our algorithm to practical graphs since the CLV model can generate graphs with power law distributed degrees, exhibited by many real-world graphs, e.g., Internet topology [26]. By Newman et al. [53], one can obtain a fairly accurate model of many social networks by using the Molloy-Reed method [50], which samples a graph from a family of random graphs with degrees distributed following a power law with exponential cutoff.

Formally, for a sequence $\mathbf{p} = (p_1, \ldots, p_n) \in [0,1]^n$, we consider the random graph $G_{\mathbf{p}}$ in which each pair of agents $i \neq j$ are friends with probability $p_i p_j$ and these events are mutually independent. This model corresponds to the setting where agents pick their edges with probabilities proportional to $\mathbf{p}$ which describes the relative distribution over the agents. Within this model, for each agent i we use the coalitional friends predictor that returns the expected number of agent i's friends within a coalition $C \subseteq [i]$, i.e., $\varphi_i(C) = p_i \sum_{j \in C} p_j$. Under this choice of predictors, we next show that the social welfare of the partition returned by MPCF *stochastically dominates* the social welfare incurred by *any* other algorithm $\mathcal{A}$, i.e., MPCF is *optimal* for graphs within the CLV model. First, we treat the case that an algorithm $\mathcal{A}$ may leave an agent i in a singleton coalition even when her neighborhood is non-empty. Specifically, we prove in Lemma 3 that an agent having no friends hinders the social welfare of the partition generated by MPCF compared to the case where his neighborhood is non-empty.

Lemma 3. *Under the CLV model, let $\mathbf{p} \in [0,1]^n$ be a weight vector. For any agent i, let $\mathbf{p}_{-i} \in [0,1]^{n-1}$ be obtained from $\mathbf{p}$ by removing its i^{th} entry. Then, for each agent $j \neq i$ and any $k \geq 0$, $\mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}})) \geq k] \leq \mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p_{-i}}})) \geq k - v_i(\mathcal{A}^\star_i(G_{\mathbf{p}}))]$ and $\mathbb{P}[v_j(\mathcal{A}^\star_j(G_{\mathbf{p}})) \geq k] \leq \mathbb{P}[v_j(\mathcal{A}^\star_j(G_{\mathbf{p_{-i}}})) \geq k - 1]$.*

Proof. Let $i \in N$. Note that an instance graph G_{-i} of $G_{\mathbf{p_{-i}}}$ can be obtained from an instance graph $G = (N, E)$ the graph $G_{\mathbf{p}}$ as follows. First, $N_{-i} := N \setminus \{i\}$ is the vertices set of the graph G_{-i}. For each $j \neq i$, recalling that F_j is agent j's set of friends (i.e., her neighborhood) in the graph G, then agent j's neighborhood in the graph G_{-i} is $F_j \setminus \{i\}$. It readily follows that the resulting graph is indeed distributed according to $G_{\mathbf{p_{-i}}}$.

We prove the result by showing that for any instance G of $G_{\mathbf{p}}$, if $\mathcal{SW}(\mathcal{A}^\star(G)) \geq k$, then $\mathcal{SW}(\mathcal{A}^\star(G_{-i})) \geq k - v_i(\mathcal{A}^\star_i(G_{\mathbf{p}}))$. Equivalently, it suffices to show that $\mathcal{SW}(\mathcal{A}^\star(G_{-i})) \geq \mathcal{SW}(\mathcal{A}^\star(G)) - v_i(\mathcal{A}^\star_i(G))$. We prove this by induction on the number of agents n. The case $n = 0$ is trivial. Thereby, we assume that $n > 0$ and that the result inductively holds for smaller values of n. Let π_t and π_t^{-i} be the partition generated by MPCF on G and G_{-i} (respectively) at time t. We proceed by the following cases:

1. If agent i resides in a singleton coalition under MPCF (i.e., $\{i\} \in \pi_n$ at the end of the execution), then she does not contribute to the social welfare of the partition π_n returned by MPCF. Since the graphs G and G_{-i} are the same except for agent i and her incident edges, we infer that $\mathcal{SW}(\mathcal{A}^\star(G)) = \mathcal{SW}(\pi_n) = \mathcal{SW}(\pi_n^{-i}) = \mathcal{SW}(\mathcal{A}^\star(G_{-i}))$.
2. Assume that agent i resides in a non-singleton coalition. Thus, let $C \in \pi_n$ be the non-singleton coalition in which agent i resides. Let $v_i(C)$ be agent i's degree in the subgraph of G induced by coalition C. Up until agent i's arrival, MPCF executed similarly on both G and G_{-i}. Specifically, $\pi_{i-1} = \pi_{i-1}^{-i}$. Upon agent i's arrival, MPCF assigns agent i to C when executed on G, i.e., $\pi_i = \{C'\}_{C \neq C' \in \pi_i^{-i}} \cup \{C \cup \{i\}\}$. The following cases are possible:
   (a) If no agent $j \in C$ has any other friend in the graph G, then MPCF will perform similarly on both G and G_{-i} from time instant i onward. Namely, $\pi_n = \{C'\}_{C \neq C' \in \pi_n^{-i}} \cup \{C \cup \{i\}\}$. Then: $\mathcal{SW}(\mathcal{A}^\star(G)) = \mathcal{SW}(\pi_n) = \mathcal{SW}(\pi_n^{-i}) + v_i(C) = \mathcal{SW}(\mathcal{A}^\star(G_{-i})) + v_i(\mathcal{A}^\star_i(G))$.
   (b) Assume that there is some agent $j \in C$ with a friend $j' > i$ in the graph G. We pick agent j' as the first friend of agent j that arrives after agent i. We distinguish between two cases:
      i. If j' has no other friend apart for j, then MPCF will incur the same social welfare at time j' for both G and G_{-i} since these graphs are the same except for agent i and her incident edges. Namely, $\mathcal{SW}(\pi_{j'}) = \mathcal{SW}(\pi_{j'}^{-i})$ as agent j' is assigned to agent j's coalition, which is of size less than α after removing agent i. Further, by the induction hypothesis: $\mathcal{SW}(\pi_n \setminus \pi_{j'}) \leq \mathcal{SW}(\pi_n^{-i} \setminus \pi_{j'}^{-i}) + v_i(C)$. We thus infer: $\mathcal{SW}(\mathcal{A}^\star(G)) = \mathcal{SW}(\pi_n) = \mathcal{SW}(\pi_{j'}) + \mathcal{SW}(\pi_n \setminus \pi_{j'}) \leq \mathcal{SW}(\pi_{j'}^{-i}) + \mathcal{SW}(\pi_n^{-i} \setminus \pi_{j'}^{-i}) + v_i(C) = \mathcal{SW}(\pi_n^{-i}) + v_i(C) = \mathcal{SW}(\mathcal{A}^\star(G_{-i})) + v_i(\mathcal{A}^\star_i(G))$.
      ii. Assume that j' has at least one other friend $j'' \neq j$. We distinguish between two cases:

A. If $j'' > j'$, then, without loss of generality, we can pick agent j'' as the next arriving friend of agent j'. Thus, MPCF will incur the same social welfare at time j' for both G and G_{-i} as these graphs are the same except for agent i and her incident edges, i.e., $\mathcal{SW}(\pi_{j'}) = \mathcal{SW}(\pi_{j'}^{-i})$. The proof proceeds in a manner identical to the previous case (case 2.b.i).

B. Assume that $j'' < j'$. Let $\tilde{C} \in \pi_{j'-1}$ such that $j'' \in \tilde{C}$. If $|\tilde{C}| = \alpha$, then agent j' cannot be assigned to $\tilde{C}$ regardless of the number friends she has within $\tilde{C}$. In that case, the proof proceeds as in case 2.b.i. Thus, we assume $|\tilde{C}| < \alpha$. If j' has more friends within $\tilde{C}$ than within C, then MPCF will assign j' to $\tilde{C}$ at time j'. If j has no friends other than j', then MPCF will perform similarly on both G and G_{-i} from time instant j'' onward and the proof proceeds as in case 2.a. If j has another friend other than j', then the proof proceeds by recursively considering the cases 2.a.i-ii.

We proved that $\mathcal{SW}(\mathcal{A}^\star(G_{-i})) \geq \mathcal{SW}(\mathcal{A}^\star(G)) - v_i(\mathcal{A}_i^\star(G_{\mathbf{p}}))$, as desired. The proof that $\mathbb{P}[v_j(\mathcal{A}_j^\star(G_{\mathbf{p}})) \geq k] \leq \mathbb{P}[v_j(\mathcal{A}_j^\star(G_{\mathbf{p_{-i}}})) \geq k-1]$ for each agent $j \neq i$ and for any $k \geq 0$ follows from similar arguments. □

Next, we show that another appealing property of MPCF is its capability of leveraging the density of a graph for attaining a higher social welfare. First, we require some terminology. Given two weight vectors $\mathbf{p}, \mathbf{p}' \in [0,1]^n$, we say that $\mathbf{p}'$ *dominates* $\mathbf{p}$ if $p_i \leq p_i'$ for every agent i. This indicates that a graph distributed as $G_{\mathbf{p}'}$ is *denser* than a graph distributed as $G_{\mathbf{p}}$. Further, we require the following:

Definition 1. (Probability for an Agent's Neighborhood) *Given some subset of agents $S \subseteq N$ and a weight vector $\mathbf{p} \in [0,1]^n$, we let $\mathcal{P}_{\mathbf{p}}^i(S)$ be the probability that the neighborhood of agent i is exactly the set S within $G_{\mathbf{p}}$. Notice that $\mathcal{P}_{\mathbf{p}}^i(S)$ solely depends on $\mathbf{p}$ and satisfies $\mathcal{P}_{\mathbf{p}}^i(S) = \prod_{j\in S} p_i p_j \prod_{j\in N\setminus S}(1-p_i p_j)$.*

In Theorem 4, we prove that the social welfare of the partition generated by MPCF for a given graph is stochastically dominated by the social welfare obtained for a *denser* one:

Theorem 4. (**MPCF's Optimality for Density Exploitation**) *Under the CLV model, let $\mathbf{p}, \mathbf{p}' \in [0,1]^n$ be s.t. $\mathbf{p}'$ dominates $\mathbf{p}$. Then, $\mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}})) \geq k] \leq \mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}'})) \geq k]$ for any $k \geq 0$.*

Proof. We only provide the proof for $\mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}})) \geq k] \leq \mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}'})) \geq k]$ for any $k \geq 0$. The proof is by induction on the number of agents n. The base case $n = 1$ is trivial since $\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}})) = \mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}'})) = 0$ with probability 1. Hence, the probabilities of reaching a social welfare of at least k is the same for both weight vectors $\mathbf{p}$ and $\mathbf{p}'$.

In the induction step, assume that $n > 0$ and the result holds for $n-1$ agents. Let $\mathbf{p}_{-n} = (p_1, \ldots, p_{n-1}) \in [0,1]^{n-1}$ be obtained from $\mathbf{p}$ by removing its last entry corresponding to agent n. Then:

$$\begin{aligned}&\mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}})) \geq k] = \mathcal{P}_{\mathbf{p}}^n(\emptyset)\mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}_{-n}})) \geq k] + \\ &+ \sum_{\emptyset\neq S\subseteq N} \mathcal{P}_{\mathbf{p}}^n(S)\mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}_{-n}})) \geq k - v_n(\mathcal{A}_n^\star(G_{\mathbf{p}}))]\end{aligned} \tag{3}$$

A similar equality is satisfied when substituting $\mathbf{p}$ by $\mathbf{p}'$.

Let $\mathbf{q} = p_n\mathbf{p}_{-n} = (p_1p_n, \ldots, p_{n-1}p_n)$ and $\mathbf{q}' = p_n'\mathbf{p}_{-n}' = (p_1'p_n', \ldots, p_{n-1}'p_n')$, which are the probabilities that agent n has an edge to agents i and i' in $\mathcal{G}_{\mathbf{p}}$ and $\mathcal{G}_{\mathbf{p}'}$ (respectively). Since $\mathbf{p}'$ dominates $\mathbf{p}$, $\mathbf{q}'$ also dominates $\mathbf{q}$. Thus, for each agent $i \in [n-1]$, we can write $1 - q_i' = (1-r_i)(1-q_i)$ for some $r_i \in [0,1]$. For each agent $i \in [n-1]$, let X_i and Y_i be independent Bernoulli variables with $\mathbf{P}[X_i = 1] = q_i$ and $\mathbf{P}[Y_i = 1] = r_i$. Further, let Z_i be the Bernoulli variable that equals 1 if either $X_i = 1$ or $Y_i = 1$, and 0 otherwise. Then, we have that $\mathbf{P}[Z_i = 1] = q_i'$. Let $M = \{i \in [n-1] : X_i = 1\}$ and $M' = \{i \in [n-1] : Z_i = 1\}$. Note that $M \subseteq M'$. For any $T \subseteq [n-1]$, the following then holds: $\mathcal{P}_{\mathbf{p}'}^n(T) = \mathbf{P}[M' = T] = \sum_{S\subseteq T}\mathbf{P}[M = S]\mathbf{P}[M' = T | M = S] = \sum_{S\subseteq T}\mathcal{P}_{\mathbf{p}}^n(S)\Delta(S,T)$, where $\Delta(S,T) := \mathbf{P}[M' = T | M = S]$. Observe that for any $S \subseteq [n-1]$, it holds that $\sum_{T\supseteq S}\Delta(S,T) = 1$. Indeed, conditioned on $M = S$, it holds that $S \subseteq M'$ with probability 1. Combined with (3), we can write $\mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}'})) \geq k]$ as:

$$\begin{aligned}&\mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}'})) \geq k] = \mathcal{P}_{\mathbf{p}'}^n(\emptyset)\mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}'_{-n}})) \geq k] + \\ &+ \sum_{\emptyset\neq T\subseteq N}\mathcal{P}_{\mathbf{p}'}^n(T)\mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}'_{-n}})) \geq k - v_n(\mathcal{A}_n^\star(G_{\mathbf{p}}))] = \\ &= \mathcal{P}_{\mathbf{p}'}^n(\emptyset)\mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}'_{-n}})) \geq k] + \sum_{\emptyset\neq T\subseteq N}\sum_{S\subseteq T}\mathcal{P}_{\mathbf{p}}^n(S)\Delta(S,T)\cdot \\ &\cdot\mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}'_{-n}})) \geq k - v_n(\mathcal{A}_n^\star(G_{\mathbf{p}}))] = \\ &= \mathcal{P}_{\mathbf{p}}^n(\emptyset)\Delta(\emptyset,\emptyset)\mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}'_{-n}})) \geq k] + \\ &+\mathcal{P}_{\mathbf{p}}^n(\emptyset)\sum_{\emptyset\neq T\subseteq N}\Delta(\emptyset,T)\mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}'_{-n}})) \geq k - v_n(\mathcal{A}_n^\star(G_{\mathbf{p}}))] + \\ &+ \sum_{\emptyset\neq S\subseteq N}\mathcal{P}_{\mathbf{p}}^n(S)\sum_{T\supseteq S}\Delta(S,T)\mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}'_{-n}})) \geq k - v_n(\mathcal{A}_n^\star(G_{\mathbf{p}}))]\end{aligned} \tag{4}$$

where the last steps follow by interchanging summations in the second term, and splitting into the cases $S = \emptyset$ and $S \neq \emptyset$.

Note that if $S \neq \emptyset$ and $S \subseteq T$, then, by MPCF, $\mathbf{p}'_{-n}$ dominates $\mathbf{p}'_n(\mathcal{A}, S)$ since the minimum weight corresponding to a neighbor in T cannot be larger than a neighbor in S. Further, since $\mathbf{p}'$ dominates $\mathbf{p}$, then $\mathbf{p}'_i(\mathcal{A}, S)$ dominates $\mathbf{p}_i(\mathcal{A}, S)$ for each agent i (up to an appropriate permutation). By the induction hypothesis, we infer that for $S \neq \emptyset$ and $T \supseteq S$ the following hold:

$$\mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}'_{-n}})) \geq k] \geq \mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}_{-n}})) \geq k] \tag{5}$$

Further, by Lemma 3 and the induction hypothesis, we obtain:

$$\begin{aligned}&\mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}'_{-n}})) \geq k - v_n(\mathcal{A}_n^\star(G_{\mathbf{p}}))] \geq \\ &\geq \mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}'_{-n}})) \geq k] \geq \mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}_{-n}})) \geq k]\end{aligned} \tag{6}$$

Applying the bounds (5)-(6) to (4), we infer:

$$\begin{aligned}&\mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}'})) \geq k] \geq \\ &\geq \mathcal{P}_{\mathbf{p}}^n(\emptyset)\Delta(\emptyset,\emptyset)\mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}_{-n}})) \geq k] + \\ &+\mathcal{P}_{\mathbf{p}}^n(\emptyset)\sum_{\emptyset\neq T\subseteq N}\Delta(\emptyset,T)\mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}_{-n}})) \geq k] + \\ &+ \sum_{\emptyset\neq S\subseteq N}\mathcal{P}_{\mathbf{p}}^n(S)\sum_{T\supseteq S}\Delta(S,T)\cdot \\ &\cdot\mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}_{-n}})) \geq k - v_n(\mathcal{A}_n^\star(G_{\mathbf{p}}))]\end{aligned} \tag{7}$$

By combining the first two terms in (7) while using since $\sum_{T\supseteq S}\Delta(S,T) = 1$ for any $S \subseteq [n-1]$ and (3), we conclude:

$$\begin{aligned}&\mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}'})) \geq k] \geq \mathcal{P}_{\mathbf{p}}^n(\emptyset)\mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}_{-n}})) \geq k] + \\ &+ \sum_{\emptyset\neq S\subseteq N}\mathcal{P}_{\mathbf{p}}^n(S)\mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}_{-n}})) \geq k - v_n(\mathcal{A}_n^\star(G_{\mathbf{p}}))] = \\ &= \mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}})) \geq k]\end{aligned} \tag{8}$$

□

We are now ready to proven our main result in Theorem 5 about MPCF's optimality under the CLV model:

Theorem 5. (**MPCF's Optimality for Welfare Maximization**) *In the CLV model, let $\mathbf{p} \in [0,1]^n$ be a weight vector and let $\mathcal{A}$ be an online algorithm for our problem. Then, for any $k \geq 0$:*

1. *For any agent i: $\mathbb{P}[v_i(\mathcal{A}_i(G_{\mathbf{p}})) \geq k] \leq \mathbb{P}[v_i(\mathcal{A}^\star_i(G_{\mathbf{p}})) \geq k]$.*
2. *$\mathbb{P}[\mathcal{SW}(\mathcal{A}(G_{\mathbf{p}})) \geq k] \leq \mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}})) \geq k]$.*

Proof. We begin with proving part (1), i.e., $\mathbb{P}[v_i(\mathcal{A}_i(G_{\mathbf{p}})) \geq k] \leq \mathbb{P}[v_i(\mathcal{A}^\star_i(G_{\mathbf{p}})) \geq k]$ for any agent i and $k \geq 0$. Let $i \in N$. The proof is by induction on the number of agents n. The base case $n = 0$ is trivial since $v_i(\mathcal{A}_i(G_{\mathbf{p}})) = v_i(\mathcal{A}^\star_i(G_{\mathbf{p}})) = 0$ with probability 1 as both algorithms will generate an empty partition. Hence, the probabilities of attaining at least k friends for agent i is the same for both algorithms (i.e., the inequality holds with equality).

For the induction step, we assume that $n > 0$ and that the result holds for $n-1$ agents. Without loss of generality, we assume that $i \neq n$. Otherwise, we could choose an agent $j \neq n$ and consider agent j instead of agent n in the sequel. Let $\mathbf{p}_{-n} \in [0,1]^{n-1}$ be obtained from $\mathbf{p}$ by removing its last entry which corresponds to agent n, i.e., $\mathbf{p}_{-n} = (p_1, \ldots, p_{n-1})$. If agents i and n are *not* friends, then n will not affect the number of agent i's friends within her assigned coalition under both algorithms, regardless of whether n joins her coalition or not. Hence, the induction hypothesis yields the required:

$$\begin{aligned} \mathbb{P}[v_i(\mathcal{A}_i(G_{\mathbf{p}})) \geq k] &= \sum_{S\subseteq N} \mathcal{P}^n_{\mathbf{p}}(S)\mathbb{P}[v_i(\mathcal{A}_i(G_{\mathbf{p}_{-n}})) \geq k] \\ &\leq \sum_{S\subseteq N} \mathcal{P}^n_{\mathbf{p}}(S)\mathbb{P}[v_i(\mathcal{A}^\star_i(G_{\mathbf{p}_{-n}})) \geq k] = \mathbb{P}[v_i(\mathcal{A}^\star_i(G_{\mathbf{p}})) \geq k] \end{aligned} \tag{9}$$

Now, assume that agents i and n are friends. If the coalition to which agent i is assigned by $\mathcal{A}$ is of size α at time $n-1$, then agent n will not affect the number of agent i's friends within her assigned coalition under $\mathcal{A}$ as agent n cannot be assigned to that coalition. Thus, the proof in that case follows from arguments similar to (9). Assume that the coalition to which agent i is assigned by $\mathcal{A}$ is of size less than α at time $n-1$. If agent n is assigned to agent i's coalition under $\mathcal{A}$, then removing agent n from the graph will decrease $v_i(\mathcal{A}_i(G_{\mathbf{p}}))$ by 1. By the induction hypothesis and Lemma 3:

$$\begin{aligned} &\mathbb{P}[v_i(\mathcal{A}_i(G_{\mathbf{p}})) \geq k] = \\ &= \sum_{S\subseteq N} \mathcal{P}^n_{\mathbf{p}}(S)\mathbb{P}[v_i(\mathcal{A}_i(G_{\mathbf{p}_{-n}})) \geq k-1] \leq \\ &\leq \sum_{S\subseteq N} \mathcal{P}^n_{\mathbf{p}}(S)\mathbb{P}[v_i(\mathcal{A}^\star_i(G_{\mathbf{p}_{-n}})) \geq k-1] \leq \\ &\leq \mathbb{P}[v_i(\mathcal{A}^\star_i(G_{\mathbf{p}})) \geq k] \cdot \sum_{S\subseteq N} \mathcal{P}^n_{\mathbf{p}}(S) = \mathbb{P}[v_i(\mathcal{A}^\star_i(G_{\mathbf{p}})) \geq k] \end{aligned} \tag{10}$$

where the last equality is due to $\sum_{S\subseteq N} \mathcal{P}^n_{\mathbf{p}}(S) = 1$.

If agent n is *not* assigned to agent i's coalition under $\mathcal{A}$, then removing agent n from the graph will not affect $v_i(\mathcal{A}_i(G_{\mathbf{p}}))$. By the induction hypothesis and Lemma 3:

$$\begin{aligned} \mathbb{P}[v_i(\mathcal{A}_i(G_{\mathbf{p}})) \geq k] &= \sum_{S\subseteq N} \mathcal{P}^n_{\mathbf{p}}(S)\mathbb{P}[v_i(\mathcal{A}_i(G_{\mathbf{p}_{-n}})) \geq k] \\ &\leq \sum_{S\subseteq N} \mathcal{P}^n_{\mathbf{p}}(S)\mathbb{P}[v_i(\mathcal{A}^\star_i(G_{\mathbf{p}_{-n}})) \geq k] \leq \\ &\leq \mathbb{P}[v_i(\mathcal{A}^\star_i(G_{\mathbf{p}})) \geq k+1] \cdot \sum_{S\subseteq N} \mathcal{P}^n_{\mathbf{p}}(S) = \\ &= \mathbb{P}[v_i(\mathcal{A}^\star_i(G_{\mathbf{p}})) \geq k+1] \leq \mathbb{P}[v_i(\mathcal{A}^\star_i(G_{\mathbf{p}})) \geq k] \end{aligned} \tag{11}$$

where we used the equality $\sum_{S\subseteq N} \mathcal{P}^n_{\mathbf{p}}(S) = 1$ and the last inequality follows from the observation that the event $\{v_i(\mathcal{A}^\star_i(G_{\mathbf{p}})) \geq k+1\}$ is a subset of the event $\{v_i(\mathcal{A}^\star_i(G_{\mathbf{p}})) \geq k\}$, and thus $\mathbb{P}[v_i(\mathcal{A}^\star_i(G_{\mathbf{p}})) \geq k+1] \leq \mathbb{P}[v_i(\mathcal{A}^\star_i(G_{\mathbf{p}})) \geq k]$.

We proceed with proving part (2), i.e., $\mathbb{P}[\mathcal{SW}(\mathcal{A}(G_{\mathbf{p}})) \geq k] \leq \mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}})) \geq k]$ for any $k \geq 0$. The proof is by induction on the number of agents n. The base case $n = 0$ is trivial since $\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}})) = \mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}})) = 0$ with probability 1 as both algorithms will generate an empty partition. Hence, the probabilities of reaching a social welfare of at least k is the same for both algorithms (i.e., the inequality holds with equality).

For the induction step, we assume that $n > 0$ and that the result holds for $n-1$ agents. Recall that $\mathbf{p}_{-n} \in [0,1]^{n-1}$ can be obtained from $\mathbf{p}$ by removing its last entry which corresponds to agent n, i.e., $\mathbf{p}_{-n} = (p_1, \ldots, p_{n-1})$. We consider the probability that algorithm $\mathcal{A}$ generates a partition with a social welfare of at least k. since $\mathbb{P}[v_i(\mathcal{A}_i(G_{\mathbf{p}})) \geq k] \leq \mathbb{P}[v_i(\mathcal{A}^\star_i(G_{\mathbf{p}})) \geq k]$ for each agent i and for any $k \geq 0$, observe that $\mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}_{-n}})) \geq k - v_n(\mathcal{A}_n(G_{\mathbf{p}}))] \leq \mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}_{-n}})) \geq k - v_n(\mathcal{A}^\star_n(G_{\mathbf{p}}))]$. By that observation and the induction hypothesis, we conclude the desired:

$$\begin{aligned} &\mathbb{P}[\mathcal{SW}(\mathcal{A}(G_{\mathbf{p}})) \geq k] = \mathcal{P}^n_{\mathbf{p}}(\emptyset)\mathbb{P}[\mathcal{SW}(\mathcal{A}(G_{\mathbf{p}_{-n}})) \geq k] + \\ &+ \sum_{\emptyset\neq S\subseteq N} \mathcal{P}^n_{\mathbf{p}}(S)\mathbb{P}[\mathcal{SW}(\mathcal{A}(G_{\mathbf{p}_{-n}})) \geq k - v_n(\mathcal{A}_n(G_{\mathbf{p}}))] \leq \\ &\leq \mathcal{P}^n_{\mathbf{p}}(\emptyset)\mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}_{-n}})) \geq k] + \\ &+ \sum_{\emptyset\neq S\subseteq N} \mathcal{P}^n_{\mathbf{p}}(S)\mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}_{-n}})) \geq k - v_n(\mathcal{A}^\star_n(G_{\mathbf{p}}))] \\ &= \mathbb{P}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}})) \geq k] \end{aligned} \tag{12}$$

□

Theorem 5 proves that MPCF is ***optimal*** for the CLV model. It also indicates that even if some online algorithm $\mathcal{A}$ satisfies Theorem 4, MPCF still exploits the graph's density better than $\mathcal{A}$ by Theorem 5. Though optimality only holds when the predictions are each agent's expected number of friends within a given coalition, MPCF is still near-optimal if the predictions are noisy due to Theorem 2.

4.2.1 The Expected Social Welfare of MPCF

We herein analyze the expected social welfare incurred by MPCF under the CLV model for $n \geq 3$, in both the asymptotic and the non-asymptotic case. Our first main result within this model is a set of equations that describe the social welfare incurred by MPCF. First, we consider the set I of the agents in $G_{\mathbf{p}}$ that are assigned to singleton coalitions by MPCF. Let $G^+_{\mathbf{p}} = (N, E^+_{\mathbf{p}})$ be the graph obtained from $G_{\mathbf{p}}$ as follows: If $p_i p_j > 0$ for a pair of agents $i \neq j$, then $(i,j) \in E^+_{\mathbf{p}}$. Let c_ℓ be the number of coalitions in $\mathcal{A}^\star(G_{\mathbf{p}})$ with exactly ℓ agents. We prove the following relation between the partition returned by MPCF and the graph $G^+_{\mathbf{p}}$:

Lemma 6. *Under the CLV model, let $\mathbf{p} \in [0,1]^n$ and let I be the agents in the graph $G_{\mathbf{p}}$ that are assigned to singleton coalitions by MPCF. Then, I is an independent set of $G^+_{\mathbf{p}}$. Further, each coalition $C \in \mathcal{A}^\star(G_{\mathbf{p}})$ is connected in $G^+_{\mathbf{p}}$ and $\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}})) \leq \sum_{\ell=1}^{\alpha} \ell(\ell-1)c_\ell$. Since $n - |I| = \sum_{\ell=2}^{\alpha} \ell c_\ell$, then $\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}})) \leq (\alpha-1)(n-|I|)$.*

Proof. Assume, towards contradiction, that the partition generated by $G_{\mathbf{p}}$ contains two isolated agents $i < j$ with $p_i p_j > 0$. Then, when j appears, MPCF adds agent j to the coalition $\{i\}$, which contradicts the fact that j is isolated. Thus, I is an independent set of $G^+_{\mathbf{p}}$. Next, note that when an agent i is inserted to an existing coalition C, there exists at least one agent $j \in C$ such that $p_i p_j > 0$ by Algorithm 1. Hence, C's social welfare is at most $|C|(|C|-1)$, yielding that $\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}})) \le \sum_{\ell=1}^{\alpha} \ell(\ell-1)c_\ell$. □

We next show how to compute $|I|$ for a *symmetric* variant of the CLV model for $n \ge 3$, where we consider the random graph $G_{\mathbf{d}}$ parameterized by the number of agents n and a vector $\mathbf{d} = \{d_i\}_{i\in[n]}$, with d_i denoting the expected degree of agent i. Any pair of agents $i \ne j$ with the same expected degree d are friends with probability d/n. As in the previous section, we can analyze MPCF when the predictions for each agent i are given by the expected number of agent i's friends within a coalition $C \subseteq N$, i.e., $\varphi_i(C) = |\{j \in C : d_j = d_i\}| \cdot d_i/n$. Let Y^d_t be the number of agents with expected degree d who are in singleton coalitions by MPCF after agent t arrives. $\{Y^d_t\}_{t\in[n]}$ is a Markov chain whose expected evolution is:

$$\mathbb{E}[Y^d_{t+1} - Y^d_t] = -\Big(1 - \Big(1 - \frac{d}{n}\Big)^{Y^d_t}\Big)\Pi_{d' \le d}\Big(1 - \frac{d'}{n}\Big)^{Y^{d'}_t} \quad (13)$$

The first term is the probability that at least one isolated agent with expected degree d is a friend of agent $t+1$, and the second term is the probability that agent $t+1$ has no isolated friend with lower expected degree (which would have been prioritized). Letting $k^d_t = -\log(1 - d/n)$ and $Z^d_t = -k^d_t \cdot Y^d_t$, (13) can be simplified as:

$$\mathbb{E}[Z^d_{t+1} - Z^d_t] = k^d_t(1 - e^{Z^d_t})\Pi_{d'<d}e^{Z^{d'}_t} \quad (14)$$

Following Kurtz [44] and many subsequent works (See, e.g., [1]), these Markov chains can be approximated by the solution of the following system of differential equations:

$$\frac{\mathrm{d}z^d(t)}{\mathrm{d}t} = k^d_t(1 - e^{z^d(t)})\Pi_{d'<d}e^{z^d(t)} \quad (15)$$

As k^d_t is independent of time t, there is a constant k^d s.t. $k^d_t \equiv k^d$. Similarly to [1, Theorem 6.1], the solution $z^d(t)$ approximates the number of isolated agents with expected degree d at time t via $-z^d(t)/k^d$. Letting $\{\delta_f\}_{f=1}^{\ell}$ be the unique expected degrees, we obtain that $|I|$ can be approximated by the quantity $\sum_{f=1}^{\ell} -z^{\delta_f}(t)/k^d$. As $\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}})) \le (\alpha-1)(n-|I|)$ by Lemma 6, we conclude that the solution to (15) thus gives an *approximate* upper bound in terms of $\{\delta_f\}_{f=1}^{\ell}$ on $\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}}))$ for the non-asymptotic case, which is exact in the asymptotic case where $n \to \infty$:

Theorem 7. (**Upper Bound on MPCF's Expected Social Welfare**) *The expected social welfare of the partition generated by MPCF when executed on $G_{\mathbf{p}}$ approaches $(\alpha-1)(n + \sum_{f=1}^{\ell} z^{\delta_f}(n)/k^{\delta_f})$ as $n \to \infty$, where $\{z^{\delta_f}(n)\}_{f=1}^{\ell}$ are the solution to the system (15).*

In [29, Appendix B], we give a similar system for the general CLV model, whose solution can be similarly derived for cases such as the symmetric CLV model, and discuss its solvability. Now, we analyze MPCF's expected social welfare for Erdős-Rényi model [35], a subclass of the CLV model where all edges in a random graph G_p independently appear with the same probability $p \in [0,1]$. Each agent's expected degree is the same and equals to $d := np$. For a wide range of the parameters n, α, p, we give in Theorem 8 an *exact* expression for MPCF's expected social welfare (up to a small additive error):

Theorem 8. (**MPCF's Expected Social Welfare for Erdős-Rényi Graphs**) *Let $p \in [0,1]$ and $n, \alpha \in \mathbb{N}$. Assume that $p = o(\log n)/n$ and $p \ge 1/n^{1+o(1)}$. Denote $M_i := v_i(\mathcal{A}^\star_i(G_p))$. Then:*

$$\mathbb{E}[M_i] = n + \alpha - \frac{\ln(e^{p\alpha} + e^{pn} - 1))}{p} \pm n^{1/2+o(1)} \quad (16)$$

In particular, $|M_i - \mathbb{E}[M_i]| = O(\sqrt{n \log n})$ with high probability dependent on n. Further, $\mathbb{E}[\mathcal{SW}(\mathcal{A}^\star(G_{\mathbf{p}}))] = \sum_{i\in[n]} \mathbb{E}[M_i] = n^2 + \alpha n - \frac{n\ln(e^{p\alpha}+e^{pn}-1))}{p} \pm n^{1.5+o(1)}$.

Proof. (*Sketch*) For any $0 \le \ell \le n-1$, we denote T_ℓ as the number of agents j such that when agent j arrives, the coalition to which agent i is assigned by MPCF has a social welfare of ℓ thus far. Further, let $\tilde{n}_i$ be minimal such that $\sum_{\ell<\tilde{n}_i} \mathbb{E}[T_\ell] \ge \alpha$. In [29, Appendix D], we first prove that $|\tilde{n}_i - \mathbb{E}[M_i]| \le n^{1/2+o(1)}$. Thus, to conclude the proof, we then show that $\tilde{n}_i$ satisfies the bound in the theorem. Specifically, we first prove that $\tilde{n}_i \ge n + \alpha - \frac{\ln(e^{p\alpha}+e^{pn}-1)}{p}$, which provides the lower bound for (16). For the upper bound, we show that $\tilde{n}_i \le n + \alpha - \frac{\ln(e^{p\alpha}+e^{pn}-1))}{p} + 2n^{o(1)}$. Hence, combining both bounds with $|\tilde{n}_i - \mathbb{E}[M_i]| \le n^{1/2+o(1)}$ yields (16). □

Remark 1. (Deterministic Graphs) *By Flammini et al. [37, Theorem 3.8], for undirected and unweighted graphs, no deterministic online algorithm has a competitive ratio better than $\alpha - 1$. For a restricted CLV model where weights are binary (i.e., $p_i \in \{0,1\}$ for any agent i), the resulting graph is deterministic and is exactly $G^+_{\mathbf{p}}$. In that case, MPCF is equivalent to the strictly α-competitive greedy algorithm devised by Flammini et al. [37, Theorem 3.9]. Therefore, we infer that MPCF on deterministic graphs is almost optimal and achieves a strict competitive ratio of α.*

5 Conclusions and Future Work

We have explored an online variant of partitioning agents in an undirected social network into coalitions of a bounded size. We gave the first results for maximizing social welfare in online hedonic games where algorithms have access to (possibly machine-learned) predictions, capturing the uncertainty regarding agents' friendships. When friendships are *uncertain*, our MPCF algorithm is *optimal* in terms of social welfare maximization for a vast family of natural random graphs. Our results can be seen as evidence that predictions are a promising tool for improving algorithms in online hedonic games, even if predictions are slightly noisy. Unexpectedly, our findings also reveal that *uncertainty* regarding agents' preferences leads to an optimal algorithm in a highly natural and common random graph model. This contrasts with scenarios *without* uncertainty, where the best known algorithm by Flammini et al. [37] is only *nearly* optimal.

Our work opens the way for many future studies that will lead to further advancements in online coalition formation. Immediate directions are exploring other classes of hedonic games in online settings, such as fractional hedonic games [10] and modified fractional hedonic games [54]. Studying other models of uncertainty and other types of predictions are also interesting directions for future research. It is also appealing to examine scenarios where assignments may be postponed, agents may be reassigned after each arrival, or both. Future research also warrants extending our work to solution concepts, such as envy-freeness [7, 15]. As our work follows a welfarist approach, another appealing direction is considering alternative welfare functions, such as the egalitarian welfare which is defined as the minimum across the agents' utilities [5].

Acknowledgements

This research was funded in part by ISF grant 1563/22.

References

[1] A. Aamand, J. Chen, and P. Indyk. (optimal) online bipartite matching with degree information. In *NeurIPS*, volume 35, pages 5724–5737, 2022.

[2] J. Alcalde and P. Revilla. Researching with whom? stability and manipulation. *Journal of Mathematical Economics*, 40(8):869–887, 2004.

[3] K. Anand, R. Ge, and D. Panigrahi. Customizing ml predictions for online algorithms. In *ICML*, pages 303–313. PMLR, 2020.

[4] A. Antoniadis, T. Gouleakis, P. Kleer, and P. Kolev. Secretary and online matching problems with machine learned advice. In *NeurIPS*, volume 33, pages 7933–7944, 2020.

[5] H. Aziz, F. Brandt, and H. G. Seedig. Optimal partitions in additively separable hedonic games. In *IJCAI*, volume 1, page 43–48, 2011.

[6] H. Aziz, F. Brandt, and H. G. Seedig. Stable partitions in additively separable hedonic games. In *AAMAS*, pages 183–190, 2011.

[7] H. Aziz, F. Brandt, and H. G. Seedig. Computing desirable partitions in additively separable hedonic games. *Artificial Intelligence*, 195:316–334, 2013.

[8] H. Aziz, S. Gaspers, J. Gudmundsson, J. Mestre, and H. Taubig. Welfare maximization in fractional hedonic games. In *IJCAI*, pages 461–467, 2015.

[9] H. Aziz, R. Savani, and H. Moulin. *Hedonic Games*, page 356–376. Cambridge University Press, 2016.

[10] H. Aziz, F. Brandl, F. Brandt, P. Harrenstein, M. Olsen, and D. Peters. Fractional hedonic games. *TEAC: ACM Transactions on Economics and Computation*, 7(2):1–29, 2019.

[11] C. Ballester. Np-completeness in hedonic games. *Games and Economic Behavior*, 49(1):1–30, 2004.

[12] S. Banerjee, H. Konishi, and T. Sönmez. Core in a simple coalition formation game. *Social Choice and Welfare*, 18(1):135–153, 2001.

[13] S. Banerjee, V. Gkatzelis, A. Gorokh, and B. Jin. Online nash social welfare maximization with predictions. In *SODA*, pages 1–19. SIAM, 2022.

[14] S. Banerjee, V. Gkatzelis, S. Hossain, B. Jin, E. Micha, and N. Shah. Proportionally fair online allocation of public goods with predictions. In *IJCAI*, pages 20–28, 2023.

[15] N. Barrot and M. Yokoo. Stable and envy-free partitions in hedonic games. In *IJCAI*, pages 67–73, 2019.

[16] N. Barrot, K. Ota, Y. Sakurai, and M. Yokoo. Unknown agents in friends oriented hedonic games: Stability and complexity. In *AAAI*, volume 33, pages 1756–1763, 2019.

[17] V. Bilò, G. Monaco, and L. Moscardelli. Hedonic games with fixed-size coalitions. In *AAAI*, volume 36, pages 9287–9295, 2022.

[18] A. Bogomolnaia and M. O. Jackson. The stability of hedonic coalition structures. *Games and Economic Behavior*, 38(2):201–230, 2002.

[19] S. Bouveret, U. Endriss, J. Lang, et al. Fair division under ordinal preferences: Computing envy-free allocations of indivisible goods. In *ECAI*, pages 387–392, 2010.

[20] J. Boyar, L. M. Favrholdt, C. Kudahl, K. S. Larsen, and J. W. Mikkelsen. Online algorithms with advice: a survey. *CSUR: ACM Computing Surveys*, 50(2):1–34, 2017.

[21] F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A. D. Procaccia. Introduction to computational social choice. *Handbook of Computational Social Choice*, pages 1–29, 2016.

[22] F. Brandt, M. Bullinger, and L. Tappe. Single-agent dynamics in additively separable hedonic games. In *AAAI*, volume 36, pages 4867–4874, 2022.

[23] M. Bullinger. Pareto-optimality in cardinal hedonic games. In *AAMAS*, pages 213–221, 2020.

[24] M. Bullinger and R. Romen. Online Coalition Formation Under Random Arrival or Coalition Dissolution. In *ESA*, volume 274 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 27:1–27:18, 2023.

[25] M. Bullinger and R. Romen. Stability in online coalition formation. In *AAAI*, volume 38, pages 9537–9545, 2024.

[26] F. Chung, L. Lu, and V. Vu. The Spectra of Random Graphs with Given Expected Degrees. *Internet Mathematics*, 1(3):257 – 275, 2003.

[27] S. Cohen and N. Agmon. Complexity of probabilistic inference in random dichotomous hedonic games. In *AAAI*, volume 37, 2023.

[28] S. Cohen and N. Agmon. Online coalitional skill formation. In *AAMAS*, pages 494–503, 2023.

[29] S. Cohen and N. Agmon. Online friends partitioning under uncertainty – supplementary materials. ECAI, https://u.cs.biu.ac.il/~agmon/CohenECAI24Sup.pdf, 2024.

[30] D. Dimitrov, P. Borm, R. Hendrickx, and S. C. Sung. Simple priorities and core stability in hedonic games. *Social Choice and Welfare*, 26(2): 421–433, 2006.

[31] Y. Dong, P. Indyk, I. Razenshteyn, and T. Wagner. Learning space partitions for nearest neighbor search. In *ICLR*, 2019.

[32] J. H. Dreze and J. Greenberg. Hedonic coalitions: Optimality and stability. *Econometrica: Journal of the Econometric Society*, pages 987–1003, 1980.

[33] E. Elkind and M. Wooldridge. Hedonic coalition nets. In *AAMAS*, volume 1, pages 417–424, 2009.

[34] E. Elkind, A. Fanelli, and M. Flammini. Price of pareto optimality in hedonic games. *Artificial Intelligence*, 288:103357, 2020.

[35] P. Erdös and A. Rényi. On random graphs. i. *Publicationes Mathematicae Debrecen*, 6:290–297, 1959.

[36] J. Feldman, A. Mehta, V. Mirrokni, and S. Muthukrishnan. Online stochastic matching: Beating 1-1/e. In *FOCS*, pages 117–126. IEEE, 2009.

[37] M. Flammini, G. Monaco, L. Moscardelli, M. Shalom, and S. Zaks. On the online coalition structure generation problem. *JAIR: Journal of Artificial Intelligence Research*, 72:1215–1250, 2021.

[38] M. Halldórsson and J. Radhakrishnan. Greed is good: Approximating independent sets in sparse and bounded-degree graphs. In *STOC*, pages 439–448, 1994.

[39] M. O. Hosli. Coalition formation and voting in the council of the european union. In *Power Indices and Coalition Formation*, pages 315–340. Springer, 2001.

[40] C.-Y. Hsu, P. Indyk, D. Katabi, and A. Vakilian. Learning-based frequency estimation algorithms. In *ICLR*, 2019.

[41] R. M. Karp and M. Sipser. Maximum matching in sparse random graphs. In *FOCS: Annual Symposium on Foundations of Computer Science*, pages 364–375. IEEE, 1981.

[42] A. M. Kerkmann, J. Lang, A. Rey, J. Rothe, H. Schadrack, and L. Schend. Hedonic games with ordinal preferences and thresholds. *JAIR: Journal of Artificial Intelligence Research*, 67:705–756, 2020.

[43] R. Kumar, M. Purohit, A. Schild, Z. Svitkina, and E. Vee. Semi-online bipartite matching. *ITCS: Innovations in Theoretical Computer Science*, 2019.

[44] T. G. Kurtz. Limit theorems for sequences of jump markov processes approximating ordinary differential processes. *Journal of Applied Probability*, 8(2):344–356, 1971.

[45] J. Lang, A. Rey, J. Rothe, H. Schadrack, and L. Schend. Representing and solving hedonic games with ordinal preferences and thresholds. In *AAMAS*, pages 1229–1237, 2015.

[46] C. Levinger, N. Hazon, S. Simola, and A. Azaria. Coalition formation with bounded coalition size. In *AAMAS*, pages 1119–1127, 2024.

[47] L. Li, E. Micha, A. Nikolov, and N. Shah. Partitioning friends fairly. In *AAAI*, volume 37, pages 5747–5754, 2023.

[48] T. Lykouris and S. Vassilvitskii. Competitive caching with machine learned advice. *JACM: Journal of the ACM*, 68(4):1–25, 2021.

[49] R. Meka, P. Jain, and I. Dhillon. In *Matrix completion from power-law distributed samples*, volume 22, 2009.

[50] M. Molloy and B. Reed. A critical point for random graphs with a given degree sequence. *Random structures & algorithms*, 6(2-3):161–180, 1995.

[51] G. Monaco, L. Moscardelli, and Y. Velaj. Stable outcomes in modified fractional hedonic games. In *AAMAS*, volume 34, pages 1–29. Springer, 2020.

[52] A. Munoz and S. Vassilvitskii. Revenue optimization with approximate bid predictions. In *NeurIPS*, volume 30, 2017.

[53] M. E. Newman, D. J. Watts, and S. H. Strogatz. Random graph models of social networks. *Proceedings of the national academy of sciences*, 99:2566–2572, 2002.

[54] M. Olsen. On defining and computing communities. In *Computing: The Australasian Theory Symposium*, volume 128, pages 97–102, 2012.

[55] G. J. Woeginger. Core stability in hedonic coalition formation. In *SOFSEM: Theory and Practice of Computer Science*, pages 33–50, Berlin, Heidelberg, 2013. Springer.