Distributed team formation in multi-agent systems:
stability and approximation

Lorenzo Coviello and Massimo Franceschetti

Abstract—We consider a scenario in which leaders are required to recruit teams of followers. Each leader cannot recruit all followers, but interaction is constrained according to a bipartite network. The objective for each leader is to reach a state of local stability in which it controls a team whose size is equal to a given constraint. We focus on distributed strategies, in which agents have only local information of the network topology and propose a distributed algorithm in which leaders and followers act according to simple local rules. The performance of the algorithm is analyzed with respect to the convergence to a stable solution.

Our results are as follows. For any network, the proposed algorithm is shown to converge to an approximate stable solution in polynomial time, namely the leaders quickly form teams in which the total number of additional followers required to satisfy all team size constraints is an arbitrarily small fraction of the entire population. In contrast, for general graphs there can be an exponential time gap between convergence to an approximate solution and to a stable solution.

I. INTRODUCTION

A multi-agent system (MAS) is composed of many interacting intelligent agents. Agents can be software, robots, or humans, and the system is highly distributed, as agents do not have a global view of the state and act autonomously of each other. These systems can be used to collectively solve problems that are difficult to solve by a single entity. Their application ranges from robotics, to disaster response, social structures, crowd-sourcing etc. A main feature of MAS is that they can manifest self-organization as well as other complex control paradigms even when the individual strategies of the agents are very simple. In short, simple local interaction can conspire to determine complex global behaviors. Examples of such emerging behaviors are in economics and game theory, where local preferences translate into global equilibria [31], in social sciences, where local exposure governs the spread of innovation [35], and in control, where local decision rules determine whether and how rapidly consensus is reached [25], [26], [32].

From a practical perspective, the performance of a MAS often depends on how quickly convergence to a global, possibly approximate, solution is reached and it is in general influenced by the network structure. For example, in the context of information diffusion in social networks, the rate of convergence of the system’s dynamics is affected by the underlaying network and the local interaction rules [18], [23].

This work was partially supported by the Army Research Office grant number W911NF-11-1-0363.

The authors are with the Department of Electrical and Computer Engineering, University of California San Diego, 9500 Gilman Dr, La Jolla CA, 92093. Emails: lcoviell@ucsd.edu, massimo@ece.ucsd.edu
from geographic constraints is given in Figure 1. In this paper, however, we consider general bipartite networks and our results do not depend on the specific constraints used to derive them.

We consider a notion of stability in which each agent controls a team of adequate size. Each leader has an incentive to reach local stability (that is, to build a team of followers of the right size) by dynamically interacting with its neighbors. The question we aim to answer is: can simple local rules lead to stable, or close to stable, team formation in reasonable time? By “close to stable” we mean that the total number of additional followers required to satisfy all team size constraints is an arbitrary small fraction of the entire population. We propose a simple, distributed, memoryless algorithm in which leaders do not communicate between each other, and we show that, in any network of size $n$, any constant approximation of a stable outcome (or of a suitably defined best outcome if a stable one does not exist) is reached in time polynomial in $n$ with high probability. In contrast, for general graphs we show through a counterexample that there can be an exponential gap between the time needed to reach stability and that needed to reach approximate stability, that is, to find the best solution compared to a good solution. We remark that, in its simplicity, the proposed algorithm is suitable to model human agents, it can be programmed on simple robots with limited computation abilities, and it is amenable to analysis.

The rest of the paper is organized as follows. After discussing how our work relates to the existing literature, in Section II we formally define the problem and the notions of stability and approximate stability, in Section III we present the distributed algorithm for leaders and followers, in Sections IV and V we present our technical results, and in Section VI we further discuss the algorithm’s performance by showing simulations’ results. To prove our result on the convergence to approximate stability, we derive a technical lemma (Lemma 1) that relates the quality of a matching to the existence of particular paths (that we call deficit-decreasing paths) of given length. The lemma extends a known combinatorial result by Hopcroft and Karp [13] to the setup of many-to-one matching, and can be considered to be of independent interest.

A. Related work

The problem of team formation that we consider is an example of distributed many-to-one matching in bipartite networks [2], [12], [30]. The one-to-one case has been previously studied in the context of theoretical computer science [20] [28]. In the control literature, our work is related to the distributed assignment problem and to group formation in MAS. In this framework, Moore and Passino [24] proposed a variant of the distributed auction algorithm for the assignment of mobile agents to tasks. Cenedese et al. [4] proposed a variant of the Stable Marriage algorithm [10] to solve the distributed task assignment problem. Abdallah and Lesser [1] proposed an “almost” distributed algorithm for coalition formation, allowing for a special agent with the role of “manager”. Gatson and den Jardins [11] studied a scenario of group formation where agents can adapt to the network structure. Tosic and Agha [33] proposed an algorithm for group formation based on the distributed computation of maximal cliques in the underlying network. Further work studied team formation in multi-robot systems [34], in the case where communication between agents is not allowed [3]. Other authors considered MAS composed by leaders and followers. Rahmani et al. [29] studied the controlled agreement problem in networks in which certain agents have leader roles, translating graph-theoretic properties into control-theoretic properties; Pasqualetti et al. [27] analyzed the problem of driving a group of mobile agents, represented by a network of leaders and followers, in which follower act according to a simple consensus rule. We distinguish ourselves from all mentioned papers, as we propose a fully distributed algorithm for group formation on arbitrary networks in which agents act according to simple local rules and perform very limited computation, and we derive performance guarantees in the form of theorems.

A more recent line of research aims to study how humans connected over a network solve tasks in a distributed fashion [6], [8], [14], [16], [17], [22]. In the work of Kearns et al. [17], human subjects positioned at the vertices of a virtual network were shown to be able to collectively reach a coloring of the network, given only local information about their neighbors. Similar papers further investigated human coordination in the case of coloring [8], [14], [22] and consensus [14], [16], with the main goal of characterizing how performance is affected by the network’s structure. Using experimental data of maximum matching games performed by human subjects in a laboratory setting, Coviello et al. [6] proposed a simple algorithmic model of human coordination that allows complexity analysis and prediction.

Finally, related to our work is also the research on social exchange networks [19], that considers a networked scenario in which each edge is associated to an economic value, nodes have to come to an agreement on how to share these values. Recently, Kanoria et al. [15] proposed a distributed algorithm that reaches approximate stability in linear time. However, we consider a different setup in which leaders have to build teams of multiple followers.

II. Problem formulation

We consider a population composed of agents of two different classes: leaders and followers. Each leader is required to recruit a team of followers whose size is equal to a given constraint, by sending requests to the followers. Followers can only accept or reject leaders’ requests. While multiple followers can be in a leader’s team, each follower can be part of a single team at a time, but is allowed to change team over time. A leader is not allowed to recruit all followers, but can only recruit the followers it is in direct communication with. The communication constraints of the population are captured by a bipartite network $G = (L \cup F, E)$ whose nodes’ partition is given by the set $L$ of leaders and the set $F$ of followers, and where there exists an edge $(f, \ell) \in E$
between follower $f$ and leader $\ell$ if and only if $f$ and $\ell$ can communicate between each other (see Figure 1). Let $N_\ell = \{f \in F : (f, \ell) \in E\}$ be the neighborhood of $\ell \in L$. For each $\ell \in L$, leader $\ell$ is required to recruit a team of $c_\ell$ followers, where $c_\ell \geq 1$.

**Definition 1 (Matching):** A subset $M \subseteq E$ is a matching of $G$ if for each $f \in F$ there exists at most a single $\ell \in L$ such that $(\ell, f) \in M$.

The definition of matching is consistent with the fact that multiple followers can be part of a leader’s team. There is a one-to-one correspondence between matchings $M$ of $G$ and tuples of teams $\{T_\ell(M) : \ell \in L\}$, where $T_\ell(M)$ denotes the team of leader $\ell$ under the matching $M$. We have that $T_\ell(M) = \{ f \in F : (\ell, f) \in M \} \subseteq N_\ell$ for every matching $M$.

We consider the following notion of stability.

**Definition 2 (Stable matching):** Given constraints $c_\ell$ for each $\ell \in L$, a matching $M$ of $G$ is stable if and only if $|T_\ell(M)| = c_\ell$ for all $\ell \in L$.

Depending on the constraints $c_\ell$, a network $G$ might not admit a stable matching. Nonetheless, given a matching of $G$, we are interested in assessing its quality. Our main result builds on the following definitions of deficit of a leader and deficit of a matching.

**Definition 3 (Deficit of a leader):** Let $\ell$ be a leader with constraint $c_\ell \geq 1$, and $M$ be a matching of $G$. The deficit of $\ell$ under the matching $M$ is $d_\ell(M) = \max(0, c_\ell - |T_\ell(M)|)$.

**Definition 4 (Deficit of a matching):** Given constraints $c_\ell \geq 1$ for each $\ell \in L$, the deficit of a matching $M$ of $G$ is

$$d(M) = \sum_{\ell \in L} d_\ell(M) = \sum_{\ell \in L} \max(0, c_\ell - |T_\ell(M)|).$$

In words, $d_\ell(M)$ is the number of additional followers leader $\ell$ needs to satisfy its size constraint. Similarly, $d(M)$ sums the numbers of additional followers each leader needs to satisfy its size constraint. Given a matching $M$, we say that a leader $\ell$ is poor if $d_\ell(M) > 0$ (that is, $|T_\ell(M)| < c_\ell$) and stable if $|T_\ell(M)| = c_\ell$. In this work, we do not consider the case of $|T_\ell(M)| > c_\ell$ since we assume that each leader $\ell$ never recruits more than $c_\ell$ followers simultaneously (since, for example, recruiting additional followers might be costly).

Observe that only poor leaders contribute to $d(M)$, and that $M$ is stable if and only if $d(M) = 0$. Given $G$, two matchings $G$ are compared with respect to their deficit, and the best matching of $G$ is defined as one minimizing the deficit.

**Definition 5 (Best matching):** A matching $M$ is a best matching of $G$ if $d(M) \leq d(M')$ for every matching $M'$.

Observe that a stable matching is also a best matching. Moreover, if $G$ admits a stable matching, $d(M)$ quantifies how much $M$ differs from a stable matching of $G$. In general, if $M'$ is a best matching of $G$ with $d(M') = d^*$, then $d(M) - d^*$ tells how much $M$ differs from a best matching of $G$.

Given a matching $M$ of $G$, the following definition provides a measure of how well $M$ approximates a best matching.

**Definition 6 (Approximate best matching):** Fix $\epsilon \in [0, 1]$, and let $m$ be the number of followers in $G$. Let $M'$ be a best matching of $G$. Then, a matching $M$ is a $(1-\epsilon)$-approximate best matching of $G$ if $d(M) - d(M') < \epsilon m$.

When $G$ admits a stable matching, we are interested in the notion of approximate stable matching.

**Definition 7 (Approximate stable matching):** Let $G$ admit a stable matching. Fix $\epsilon \in [0, 1]$, and let $m$ be the number of followers in $G$. Then, a matching $M$ is a $(1-\epsilon)$-approximate stable matching of $G$ if $d(M) < \epsilon m$.

III. The Algorithm

We now present a distributed algorithm for team formation. Time is divided into rounds, and each round is composed by two stages. In the first stage, each leader acts according to the algorithm in Table 1, and in the second stage each follower acts according to the algorithm in Table 2.

First consider a leader $\ell$, and let $M$ be the matching at the beginning of a given round. If $\ell$ is poor (that is, $|T_\ell(M)| < c_\ell$) and $|T_\ell(M)| < |N_\ell|$ (that is, $\ell$ is not already matched with all followers in $N_\ell$) then, with probability $p$ (where $p \in (0, 1]$ is a fixed constant), $\ell$ attempts to recruit an additional follower, chosen as explained below, by sending a matching request. An unmatched follower in $N_\ell$, if any, is chosen uniformly at random; otherwise, a follower in $N_\ell \setminus T_\ell(M)$ is chosen uniformly at random. In other words, leaders always prefer to recruit followers that are currently unmatched over matched ones. Note that a leader tries to recruit an additional follower after checking if local stability holds (that is, after checking if its team size is equal to $c_\ell$).

Consider now a follower $f$. During each round, if $f$ has incoming requests then each request is rejected independently of the others with probability $1-\eta$ (where $\eta \in (0, 1]$ is a fixed constant). If all incoming requests are rejected, then $f$ does not change team (if currently matched) or it remains unmatched (if currently unmatched). Otherwise, one among the active requests is chosen uniformly at random, $f$ joins the corresponding leader, and all the other requests are discarded. For ease of presentation, we assume that a follower is equally likely to join a team when unmatched and to change team when currently matched, but all our results hold in general.

<table>
<thead>
<tr>
<th>Table 1 Algorithm for leader $\ell \in L$</th>
</tr>
</thead>
<tbody>
<tr>
<td>if $</td>
</tr>
<tr>
<td>then with probability $p$ do the following</td>
</tr>
<tr>
<td>if $\exists$ unmatched $f \in N_\ell$ then</td>
</tr>
<tr>
<td>choose an unmatched follower $f' \in N_\ell$ u.a.r.</td>
</tr>
<tr>
<td>else</td>
</tr>
<tr>
<td>choose a follower $f' \in N_\ell \setminus T_\ell(M)$ u.a.r.</td>
</tr>
<tr>
<td>end if</td>
</tr>
<tr>
<td>send a matching request to $f'$</td>
</tr>
<tr>
<td>end if</td>
</tr>
</tbody>
</table>

The proposed algorithm is memoryless, the actions of each agent only depend on local information, and the leaders do not communicate between each other. Also, it is self-stabilizing, that is, once a stable matching is reached, leaders stop recruiting followers. Moreover, it is a single-stage algorithm, that is, agents never change their behavior until stability is reached. Finally, observe that the exchanged messages can be represented by a single bit.
Table 2 Algorithm for follower $f \in F$

<table>
<thead>
<tr>
<th>Algorithm for follower $f \in F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>if $f$ has incoming requests then</td>
</tr>
<tr>
<td>for each leader $\ell$ requesting $f$ do</td>
</tr>
<tr>
<td>with probability $1 - q$ reject $\ell$'s request</td>
</tr>
<tr>
<td>end for</td>
</tr>
<tr>
<td>if there are active requests then</td>
</tr>
<tr>
<td>select one u.a.r. and join the corresponding team</td>
</tr>
<tr>
<td>reject all other requests</td>
</tr>
<tr>
<td>end if</td>
</tr>
<tr>
<td>end</td>
</tr>
</tbody>
</table>

IV. CONVERGENCE TO APPROXIMATE STABLE MATCHINGS

In this section, we only consider networks admitting stable matchings, and we show that, given any network and any constant $\varepsilon \in (0, 1)$, a $(1 - \varepsilon)$-approximate stable matching is reached in a number of rounds that is polynomial in the network size with high probability. All our results also hold for reaching approximate best matchings, by replacing $d(M)$ with $d(M) - d(M')$, where $M'$ is a best matching of $G$.

Given a network $G$, for every $t \geq 0$, let $M(t)$ be the matching at the beginning of round $t$, with deficit $d(M(t))$. The next properties follow from the fact that leaders do not voluntarily disengage from the followers (and the deficit of a leader increases of a unit only if the deficit of another one decreases by one unit), and the assumption $c_i \geq 1, \forall \ell$.

Property 1: For $t \geq 0$, $d(M(t))$ is non-increasing in $t$.

Property 2: $d(M(t)) \leq m$ for every $t \geq 0$.

We are now ready to state our main result. The maximum degree of the leaders is given by $\Delta = \max_{\ell \in L} |N(\ell)|$.

Theorem 1: Let $G$ be a network with $m$ followers and which admits a stable matching. Let $\Delta$ be the maximum degree of the leaders. Fix $0 < \varepsilon < 1$, and let $c(\varepsilon) = \lfloor 1/\varepsilon \lfloor (pq)^{-1/\varepsilon}$ and $a(\varepsilon) = \lfloor 1/\varepsilon \rfloor$. Then, a $(1 - \varepsilon)$-approximate stable matching of $G$ is reached within $c(\varepsilon) \Delta a(\varepsilon) m^2$ rounds with high probability.

Example 1: If $\Delta$ is constant in the network size, then one can choose $\varepsilon = 1/\log m$, and Theorem 1 implies that a $(1 - 1/\log m)$-approximate stable matching is reached in at most $m^2 \log(\Delta/pq) \log m$ rounds with high probability.

To prove Theorem 1, we introduce the notion of deficit-decreasing path, that, in our setup, plays the role of the augmenting path in the context of one-to-one matching [7].

Note that a path alternates leaders and followers.

Definition 8 (Deficit-decreasing path): Given a matching $M$ of $G$, a cycle-free path $P = \ell_0, f_1, \ell_1, \ldots, f_k$ (of odd length 2k-1) is a deficit-decreasing path relative to $M$ if $(\ell_i, f_i) \in M$ for all $1 \leq i \leq k - 1$, $\ell_0$ is a poor leader, and $f_k$ is an unmatched follower.

In words, a deficit-decreasing path starts at a poor leader with an edge not in $M$, ends at a follower that is not matched, and alternates edges in $M$ and edges not in $M$. To justify the nomenclature, observe that, if $d(M) > 0$ and $P$ is a deficit-decreasing path relative to $M$, a new matching $M'$ such that $d(M') = d(M) - 1$ can be obtained by turning each unmatched edge of $P$ into a matched edge, and vice versa (see Fig. 2).

The proof of Theorem 1 builds on a technical result that, given a matching $M$ with $d(M) \geq em$, guarantees the existence of a deficit-decreasing path of length at most $2[1/\varepsilon^2]$. The existence of such a path allows us to bound the number of rounds needed for a one-unit reduction of the deficit. Our technical lemma, proven in [5] extends a known result by Hopcroft and Karp [13, Theorem 1] given in the context of one-to-one matchings, but our proof is more subtle because leaders can be matched to multiple followers and can have different size constraints $c_i$. The symmetric difference of two sets $A$ and $B$ is $A \oplus B = (A \setminus B) \cup (B \setminus A)$. Two paths are follower-disjoint if they do not share any follower.

Lemma 1: Let $G$ admit a stable matching $N$. Let $M$ be a matching of $G$ with deficit $d(M) > 0$. Then, in $M \oplus N$ there are at least $d(M)$ follower-disjoint deficit-decreasing paths relative to $M$.

We make use of Lemma 1 through the following corollary.

Corollary 1: Let $G$ be a network with $m$ followers, admitting a stable matching $N$. Let $M$ be a matching of $G$ with deficit $d(M) \geq em$, for some $\varepsilon > 0$. Then, in $M \oplus N$ there exists a deficit-decreasing path relative to $M$ of length at most $2[1/\varepsilon^2]$.

Proof: By Lemma 1, if $d(M) \geq em$ and $N$ is a stable matching of $G$, then in $M \oplus N$ we can choose $em$ follower-disjoint deficit-decreasing paths relative to $M$ of cumulative length at most $2m$ (as they do not share followers, and $G$ is bipartite). One of them has length at most $2[1/\varepsilon^2]$. $\blacksquare$

A. Proof of Theorem 1

Let $G$ be a network with $n$ leaders and $m$ followers, admitting a stable matching. Fix $0 < \varepsilon < 1$. Recall that we denote the matching at the beginning of round $t \geq 0$ by $M(t)$. For every $0 < x \leq 1$, let

$$\tau(x) = \min \{ t : d(M(t)) < xm \}$$

be the first round at which the deficit becomes strictly smaller than $xm$. By Property 1, $\tau(x_2) \geq \tau(x_1)$ if $x_2 < x_1$. We are interested in bounding $\tau(\varepsilon^2)$.

Consider any round $t_1 \geq 0$. By Property 2, $d(M(t_1)) \leq m$, and therefore there exists $0 < \varepsilon' \leq 1$ such that $d(M(t_1)) = \varepsilon'm$ (we assume $\varepsilon' > 0$, as the case of $\varepsilon' = 0$ is trivial). Observe that $\tau(\varepsilon')$ can be equivalently defined as

$$\tau(\varepsilon') = \min \{ t > t_1 : d(M(t)) < d(M(t_1)) \},$$

that is, the first round after $t_1$ in which the deficit decreases. The following lemma, proven in [5] bounds $\tau(\varepsilon')$. 2758
Lemma 2: Let $\Delta$ be the maximum degree of the leaders in $G$. Let $d(M(t_1)) = e'm$, for some $0 < e' \leq 1$. Let $c(e') = \lfloor 1/e' \rfloor$ and $\alpha(e') = \lfloor 1/e' \rfloor$. Then

$$P(e' - t_1 \leq c(e')\Delta^{a(e')}m) \geq 1 - e^{-m/8}.$$  

We use the result of Lemma 2 to prove the claim of the theorem. If the matching $M(0)$ at time $t = 0$ is not stable then there exists $0 < e_0 \leq 1$ such that $d(M(0)) = e_0m$, and therefore the number of rounds $\tau(e_0)$ before a one-unit reduction of the deficit is at most $c(e_0)\Delta^{a(e_0)}m$ with probability at least $1 - e^{-m/8}$. Similarly, if $M(\tau(e_0))$ is not stable, then $d(M(e_0)) = e_1m$ for some $e_1 < e_0$. Iterating the same argument, the deficit decreases by one unit in at most $c(e_i)\Delta^{a(e_i)}m$ rounds with probability at least $1 - e^{-m/8}$, reaching a matching $M(\tau(e_i))$ with $d(M(e_i)) = e_2m < e_1m$. Given the target $e$, as $e_0 \leq 1$ and the deficit decreases by integer amounts, then there exist $k \leq e_0m \leq m$ and a sequence $1 \geq e_0 > \cdots > e_{k-1} > e_0 > 0$ such that $e_{k-1} \geq e > e_k$. By union bound, with probability at least $1 - me^{-m/8}$.

$$\tau(e) \leq \sum_{i=0}^{k-1} c(e_i)\Delta^{a(e_i)}m < c(e_{k-1})\Delta^{a(e_{k-1})}m^2 < c(e)\Delta^{a(e)}m^2.$$  

V. Exponential convergence

Theorem 1 gives a polynomial bound for reaching a $(1-\varepsilon)$-approximate stable matching for any constant $0 < \varepsilon < 1$ and any network. However, a similar guarantee cannot be derived for the case of a stable matching, as shown in this section through a counterexample. In particular, we define a sequence of networks of increasing size and maximum degree that diverges with the network size, and show that the number of rounds required to converge from an approximate matching to the stable matching is exponentially large in the network’s size with high probability from an overwhelming fraction of the approximate matchings $M$ such that $d(M) = 1$.

For $n \geq 1$, let $G_n = (L_n \cup F_n, E_n)$ be the network with $n$ leaders and $n$ followers (i.e., $L_n = \{\ell_1, \ldots, \ell_n\}$ and $F_n = \{f_1, \ldots, f_n\}$), with edges $E_n = \{(\ell_i, f_i) : 1 \leq i \leq n, j \leq l\}$, and team size constraints $c_\ell = 1$ for all $\ell \in L_n$, see Figure 3. $G_n$ has maximum degree $\Delta = n$ and a unique stable matching given by $M_n^* = \{(\ell_i, f_i) : 1 \leq i \leq n\}$.

Here we only provide a sketch of the proof, whose details are presented in [5]. To get an understanding of the algorithm’s dynamics, consider the matching $M_n = \{(\ell_i, f_{i-1}) : 2 \leq i \leq n\}$, highlighted in Figure 3 for the case of $n = 6$. Observe that $d(M_n) = 1$ and, under $M_n$, $\ell_1$ is poor, and the remaining leaders are stable. According to the algorithm, $\ell_1$ attempts to recruit $f_1$ (currently in $\ell_2$’s team). If $f_1$ accepts, then $\ell_1$ becomes stable and $\ell_2$ becomes poor (and can in turn attempt to recruit either $f_1$ or $f_2$). After each round, there is a unique poor leader, until the stable matching is reached.

In general, fix any matching $M$ of $G_n$ such that $d(M) = 1$. In $M$, there is a single poor leader $\ell_0$, and a single unmatched follower $f_k$. $M$ is associated to a unique deficit-decreasing path $\ell_{i_0}, f_{i_0}, \ldots, \ell_{i_k-1}, f_{i_k}, \ell_{i_k}$, $f_k$. If $k \geq 1$, we define $h(M) = i_{k-1}$ as the height of $M$.

Starting from $M$, for every $t < \tau(M)$, the matching $M(t)$ at round $t$ has deficit $d(M(t)) = 1$ (by Property 1), a single poor leader $\ell_{i(t)}$, the single unmatched follower $f_{i_k}$ and height $h(M(t)) = h(M) = i_{k-1}$. The stochastic process tracking the position of the poor leader $\ell_{i(t)}$ is not a classical random walk on $\{\ell_1, \ldots, \ell_n\}$ and its transition probabilities at each round depend on the current matching. The time to reach stability is upper bounded by $\min\{t : \tau(t) = h(M)\}$, that is, the first round in which $\ell_{i(t)}$ becomes poor.

We then consider a one-to-one correspondence between the matchings $M(t)$ reachable from $M$ and the nodes of a tree whose size is exponentially large in the height $h(M)$. In particular, we show that the process $\{M(t) : t \geq 0, M(0) = M\}$ is equivalent to a random walk on the nodes of the tree, and that reaching the matching with $i(t) = h(M)$ corresponds to reaching the root of the tree. A random walk starting at any node of the tree visits the root after a number of steps that is exponentially large in the height $h(M)$ with high probability.

The proof of Theorem 2 is completed by arguing that, for any constant $0 < \gamma < 1$, a $1 - 2^{-2(1-\gamma)n}$ fraction of all matchings $M$ of $G_n$ such that $d(M) = 1$ have height $h(M) \geq \gamma n$.

VI. Simulations

In Figure 4, the algorithm’s average convergence time on the sequence of networks $G_n$ defined in Section V is shown (in logarithmic scale). On the one hand, the thick solid line suggests that the average number of rounds to reach a 0.9-approximate stable matching is upper bounded by a polynomial of small degree, consistently with Theorem 1. On the other hand, convergence to the stable matching requires an average number of rounds that grows exponentially in $n$ (thin solid line), as predicted by Theorem 2. In addition, the dotted line represents the average time after which all followers become matched.

Figure 5 shows the algorithm’s performance in reaching successively finer approximations of the best matching on random networks $G(n,m,\rho)$. Here, $G(n,m,\rho)$ refers to a random bipartite network with $n$ leaders and $m$ followers, in which each edge exists independently of the others with probability $\rho$ (we fixed $\rho = 0.04$), and with constraint $c_\ell = \min\{m/n,|N|\}$ for each leader $\ell$. For each of the $(n,m)$ pairs that we considered, 20 random $G(n,m,\rho)$ were
generated, and the algorithm was run 20 times on each. The plot suggests that a good solution is reached quickly, while most of the time is spent improving it to the best solution.

VII. DISCUSSION

The distributed algorithm we proposed, in which leaders and followers act according to simple local rules, is computationally tractable and allows us to derive performance guarantees in the form of theorems. Despite its simplicity, the algorithm is shown to reach an arbitrarily close approximation of a stable (or a best) matching in polynomial time in any network. In general there can be an exponential gap between reaching an approximate solution and a stable solution.

In the proposed algorithm, leaders do not communicate between each other, and only act in response to their own status and the status of their neighborhoods. How communication between leaders affects performance is an open question, as well as determining what amounts of communication and complexity are necessary to remove the exponential gap in the case of unbounded degree networks.

REFERENCES