



How long does it take for all users in a social network to choose their communities?

Jean-Claude Bermond, Augustin Chaintreau, Guillaume Ducoffe, Dorian Mazaure

► To cite this version:

Jean-Claude Bermond, Augustin Chaintreau, Guillaume Ducoffe, Dorian Mazaure. How long does it take for all users in a social network to choose their communities?. FUN 2018 - 9th International Conference on Fun with Algorithms, 2018, La Maddalena, Italy. hal-01780627

HAL Id: hal-01780627

<https://inria.hal.science/hal-01780627>

Submitted on 27 Apr 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

How long does it take for all users in a social network to choose their communities?

Jean-Claude Bermond

Université Côte d’Azur, CNRS, Inria, I3S, France

Augustin Chaintreau

Columbia University in the City of New York

Guillaume Ducoffe¹

National Institute for Research and Development in Informatics and Research Institute of the University of Bucharest, București, România

Dorian Mazauric²

Université Côte d’Azur, Inria, France

Abstract

We consider a community formation problem in social networks, where the users are either friends or enemies. The users are partitioned into conflict-free groups (*i.e.*, independent sets in the *conflict graph* $G^- = (V, E)$ that represents the enmities between users). The dynamics goes on as long as there exists any set of at most k users, k being any fixed parameter, that can change their current groups in the partition *simultaneously*, in such a way that they all strictly increase their utilities (number of friends *i.e.*, the cardinality of their respective groups minus one). Previously, the best-known upper-bounds on the maximum time of convergence were $\mathcal{O}(|V|\alpha(G^-))$ for $k \leq 2$ and $\mathcal{O}(|V|^3)$ for $k = 3$, with $\alpha(G^-)$ being the independence number of G^- . Our first contribution in this paper consists in reinterpreting the initial problem as the study of a dominance ordering over the vectors of integer partitions. With this approach, we obtain for $k \leq 2$ the tight upper-bound $\mathcal{O}(|V| \min\{\alpha(G^-), \sqrt{|V|}\})$ and, when G^- is the empty graph, the exact value of order $\frac{(2|V|)^{3/2}}{3}$. The time of convergence, for any fixed $k \geq 4$, was conjectured to be polynomial [7, 14]. In this paper we disprove this. Specifically, we prove that for any $k \geq 4$, the maximum time of convergence is an $\Omega(|V|^{\Theta(\log |V|)})$.

2012 ACM Subject Classification Networks, Theory of computation

Keywords and phrases communities, social networks, integer partitions, coloring games, graphs

Digital Object Identifier 10.4230/LIPIcs...

Foreword: The organizers of a wedding (party) have difficulties in arranging place settings for the guests as there are many incompatibilities among those who do not want to be at the same table as an "enemy" (ex girl (boy) friend, boss or employee, student or supervisor, etc...). The organizers realize that they have no set of 5 pairwise friends and so allow people place themselves. Successively each person joins a table where she has no enemies or starts a new table. At any time a person can move from one table to another table (of course where she has no enemy) if in doing so she increases strictly the number of

¹ Part of this work has been done as PhD student in the project Coati at Université Côte d’Azur and during visits at Columbia University in the City of New York. This work was also supported by the Institutional research programme PN 1819 "Advanced IT resources to support digital transformation processes in the economy and society - RESINFO-TD" (2018), project PN 1819-01-01 "Modeling, simulation, optimization of complex systems and decision support in new areas of IT&C research", funded by the Ministry of Research and Innovation, Romania.

² Part of this work has been done during his post-doc at Columbia University in the City of New York.



37 friends she has at the new table. The process converges relatively fast (linear time). Some
 38 time later the organizers of FUN having heard about this scenario decide to use the same
 39 process to place the participants in different groups for the social activities of the afternoon.
 40 Each participant registers first in her own group. The organizers decide to accelerate the
 41 process by authorizing not just one person but any subset of 4 persons to change their mind
 42 and leave the group in which they are registered to join another group or create a new group;
 43 these persons move only if they desire to do so, that is, they increase strictly the number
 44 of friends. Surprisingly the process takes a very long (exponential) time and night arrives
 45 before groups are formed. As we will discover, the exponential time derives from the fact
 46 that at FUN all the persons are friends and there are no enemies due to the use of moves
 47 implying 4 persons. At this point the reader (and the organizers) might ask why we see such
 48 a difference in behaviors and how long does it takes for users of a social network to form
 49 groups. The answers to these questions and "all you wanted to know but were afraid to ask"
 50 will be revealed in this paper.

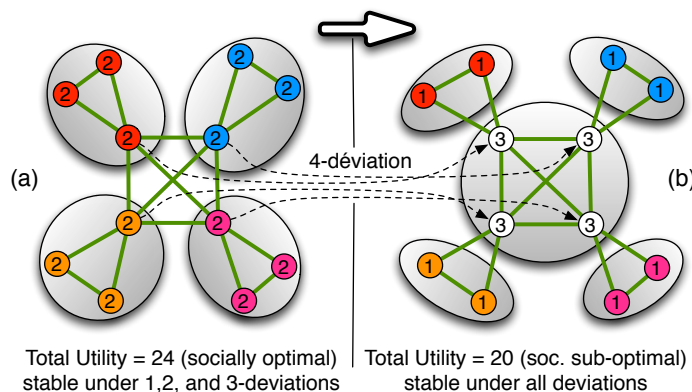
51 **1 Introduction**

52 Community formation is a fundamental problem in social network analysis. It has already
 53 been modeled in several ways, each trying to capture key aspects of the problem. The model
 54 studied in this paper has been proposed in [14] in order to reflect the impact of information
 55 sharing on the community formation process. Although it is a simplified model, we show that
 56 its understanding requires us to solve combinatorial problems that are surprisingly intricate.
 57 More precisely, we consider the following dynamics of formation of groups (communities)
 58 in social networks. Each group represents a set of users sharing about some information
 59 topic. We assume for simplicity that each user shares about a given topic in only one group.
 60 Therefore the groups will partition the set of users. We follow the approach of [14]. An
 61 important feature is the emphasis on incompatibility between some pairs of users that we
 62 will call enemies. Two enemies do not want to share information and so will necessarily
 63 belong to different groups. In the general model one consider different degrees of friendship
 64 or incompatibilities. Here we will restrict to the case where two users are either friends
 65 or enemies – as noted in [14], even a little beyond this case, the problem quickly becomes
 66 intractable. As example, if we add a neutral (indifference) relation, there are instances for
 67 which there is no stability.

68 The social network is often modeled by the friendship graph G^+ where the vertices are
 69 the users and an edge represents a friendship relation. We will use this graph to present the
 70 first notions and examples. However, for the rest of the article and the proofs we will use the
 71 complementary graph, that we call the *conflict graph* and denote by G^- ; here the vertices
 72 represent users and the edges represent the incompatibility relation. We assign each user a
 73 *utility* which is the number of friends in the group to which she belongs. Equivalently, the
 74 utility is the size of the group minus one, as in a group there is no pair of enemies; in [14]
 75 this is modeled by putting the utility as $-\infty$ when there is an enemy in the group.

76 In the example of Figure 1, the graph depicted is the friendship graph: the edges represent
 77 the friendship relation, and if there is no edge it corresponds to a pair of enemies. Figure 1(a)
 78 depicts a partition of 12 users composed of 4 non-empty groups each of size 3. The integers
 79 on the vertices represent the utilities of the users which are all equal to 2. Figure 1(b)
 80 depicts another partition consisting of 5 groups with one group of size 4 (where users have
 81 utility 3) and 4 groups of size 2 (where users have utility 1).

82 In this study we are interested in the dynamics of formation of groups. Another important



■ **Figure 1** A friendship graph with 12 vertices (users). (a) 3-stable partition that is not 4-stable but it is optimal in terms of total utility. (b) k -stable partition for any $k \geq 1$ that is not optimal in terms of total utility.

feature of [14], taken into account in the dynamics, is the notion of bounded cooperation between users. More precisely, the dynamics is as follows: initially each user is alone in her own group. In the simplest case, a move consists for a specific user to leave the group to which she belongs to join another group but only if this action increases strictly her utility (acting in a selfish manner); in particular, it implies that a user does not join a group where she has an enemy. In the k -bounded mode of cooperation, a set of at most k -users can leave their respective groups to join another group, again, only if each user increases strictly their utility. If the group they join is empty it corresponds to creating a new group. We call such a move a k -deviation. Note that this notion is slightly different from that of $(k+1)$ -defection of [14]. We will say that a partition is k -stable if there does not exist a k -deviation for this partition.

The partition of Figure 1(a) is k -stable when $k \in \{1, 2, 3\}$. Indeed each user has at least one enemy in each non empty other group and so cannot join another group. Furthermore, when $k \leq 3$, if k users join an empty group their utility will be at most 2 and so will not strictly increase. However, this partition is not 4-stable because there is a 4-deviation: the four central users can join an empty group and so they increase their utilities from 2 to 3. The partition obtained after such a 4-deviation is depicted in Figure 1(b). This partition is k -stable for any $k \geq 1$. Note that the utility of the other users is now 1 (instead of 2). Thus, we deduce that this partition is not optimal in terms of total utility (the total utility has decreased from 24 to 20); but it is now stable under all deviations. This illustrates the fact that users act in a selfish manner as some increase their utility, but on the contrary the global utility decreases. For more information on the suboptimality of k -stable partitions, *i.e.*, bounds on the price of anarchy and the price of stability, the reader is referred to [14].

1.1 Related work.

This above dynamics has been also modeled in the literature with *coloring games*. A coloring game is played on the conflict graph. Players must choose a color in order to construct a proper coloring of the graph, and the individual goal of each agent is to maximize the number of agents with the same color as she has. On a more theoretical side, coloring games have been introduced in [18] as a game-theoretic setting for studying the chromatic number in graphs. Specifically, the authors in [18] have shown that for every coloring game, there

exists a Nash equilibrium where the number of colors is exactly the chromatic number of the graph. Since then, these games have been used many times, attracting attention in the study of information sharing and propagation in graphs [4, 7, 14]. Coloring games are an important subclass of the more general Hedonic games, of which several variations have been studied in the literature in order to model coalition formation under selfish preferences of the agents [10, 12, 15, 5, 8, 16]. We stress that while every coloring game has a Nash equilibrium that can be computed in polynomial-time [18], deciding whether a given Hedonic game admits a Nash equilibrium is NP-complete [1].

If the set of edges of the conflict graph is empty (edgeless conflict graph), there exists a unique k -stable partition namely that consists of the group of all the users. In [14], it is proved that there always exists a k -stable partition for any conflict graph, but that it is NP-hard to compute one (this result was also proved independently in [7]). Indeed, if k is equal to the number of users, a largest group in such a partition must be a maximum independent set of the conflict graph. In contrast, it can be computed a k -stable partition in polynomial time for every fixed $k \leq 3$, by using simple *better-response dynamics* [18, 7, 14]. In such an algorithm one does a k -deviation until there does not exist any one. That corresponds to the dynamics of formation of groups that we study in this work for larger values of k .

1.2 Additional related work and our results.

In this paper we are interested in analyzing in this simple model the convergence of the dynamics with k -deviations, in particular in the worst case. It has been proved implicitly in [14] that the dynamics always converges within at most $\mathcal{O}(2^n)$ steps. Let $L(k, G^-)$ be the size of a longest sequence of k -deviations on a conflict graph G^- . We first observe that the maximum value, denoted $L(k, n)$, of $L(k, G^-)$ over all the graphs with n vertices is attained on the edgeless conflict graph G^0 of order n . Prior to this work, no lower bound on $L(k, n)$ was known, and the analysis was limited to potential function that only applies when $k \leq 3$ [7, 14] giving upper bounds of $\mathcal{O}(n^2)$ in the case $k = 1, 2$ and $\mathcal{O}(n^3)$ in the case $k = 3$. In order to go further in our analysis, the key observation is that when the conflict graph is edgeless, the dynamics depends only of the size of the groups of the partitions generated. Following [3], let an integer partition of $n \geq 1$, be a non-increasing sequence of integers $Q = (q_1, q_2, \dots, q_n)$ such that $q_1 \geq q_2 \geq \dots \geq q_n \geq 0$ and $\sum_{i=1}^n q_i = n$. If we rank the groups by non increasing order of their size, there is a natural relation between partition in groups and integer partitions (the size q_i of the group X_i corresponding to the integers q_i of the partition of n). Using this relation, we prove in Section 3 that the better response dynamics algorithm reaches a stable partition in p_n steps, where $p_n = \Theta((e^{\pi\sqrt{2n/3}})/n)$ denotes the number of integer partitions. This is already far less than 2^n , which was shown to be the best upper bound that one can obtain for $k \geq 4$ when using an additive potential function [14].

Table 1 summarizes our contributions described below.

- For $k = 1, 2$, we refine the relation between partitions into groups and integer partitions as follows.
 - In the case $k = 1$ (Section 4.1), we prove that there is a one to one mapping between sequences of 1-deviations in the edgeless conflict graph and chains in the dominance lattice of integer partitions. Then, we use the value of the longest chain in this dominance lattice obtained in [9] to determine exactly $L(1, n)$. More precisely, if $n = \frac{m(m+1)}{2} + r$, with $0 \leq r \leq m$, $L(1, n) = 2^{\binom{m+1}{3}} + mr$. The latter implies in particular $L(1, n)$ is of order $\mathcal{O}(n^{\frac{3}{2}})$, thereby improving the previous bound $\mathcal{O}(n^2)$.

k	Prior to our work	Our results	
1	$\mathcal{O}(n^2)$ [14]	exact analysis, which implies $L(1, n) \sim \frac{(2n)^{3/2}}{3}$	Theorem 6
2	$\mathcal{O}(n^2)$ [14]	exact analysis, which implies $L(2, n) \sim \frac{(2n)^{3/2}}{3}$	Theorem 9
1-2	$\mathcal{O}(n\alpha(G^-))$ [18]	$L(k, G^-) = \Omega(n\alpha(G^-))$ for some G^- and $\alpha(G^-) = \mathcal{O}(\sqrt{n})$	Theorem 12
3	$\mathcal{O}(n^3)$ [7, 14]	$L(3, n) = \Omega(n^2)$	Theorem 13
≥ 4	$\mathcal{O}(2^n)$ [14]	$L(k, n) = \Omega(n^{\Theta(\ln(n))})$, $L(k, n) = \mathcal{O}(\exp(\pi\sqrt{2n/3})/n)$	Theorem 14

■ **Table 1** Previous bounds and results we obtained on $L(k, n)$ and $L(k, G^-)$.

- 159 ■ In Section 4.2, we prove that any 2-deviation can be “replaced” (in some precise way)
 160 either by one or two 1-deviations, and so, $L(2, n) = L(1, n)$.
 161 ■ For $k = 1, 2$ and a general conflict graph G^- , the value of $L(k, G^-)$ depends on the
 162 independence number $\alpha(G^-)$ (cardinality of a largest independent set) of the conflict
 163 graph. In [18] it was proved that the convergence of the dynamics is in $\mathcal{O}(n\alpha(G^-))$.
 164 In the case of edgeless conflict graph, we have seen that $L(1, n) = \Omega(n^{3/2})$ and so
 165 the preceding upper-bound was not tight. So we inferred that the convergence of the
 166 dynamics was in $\mathcal{O}(n\sqrt{\alpha(G^-)})$. Yet in fact we prove in Section 4.3 that, for any
 167 $\alpha(G^-) = \mathcal{O}(\sqrt{n})$, there exists a conflict graph G^- with n vertices and independence
 168 number $\alpha(G^-)$ for which we need a sequence of at least $\Omega(n\alpha(G^-))$ 1-deviations to
 169 reach a stable partition. *For the wedding’s example of the foreword, $\alpha(G^-) = 4$ and*
 170 *so the sequence is linearly bounded.*
 171 ■ Finally, our main contribution is obtained for $k \geq 3$. Prior to our work, it was known
 172 that $L(3, n) = \mathcal{O}(n^3)$, which follows from another application of the potential function
 173 method [14]. But nothing proved that $L(3, n) > L(2, n)$, and in fact it was conjectured
 174 in [7] that both values are equal. In Section 5.2, we prove (Theorem 13) that $L(3, n) =$
 175 $\Omega(n^2)$ and thus we show for the first time that deviations can delay convergence and that
 176 the gap between $k = 2$ and $k = 3$ obtained from potential function is indeed justified. It
 177 was also conjectured in [14] that $L(k, n)$ was polynomial in n for k fixed. In Section 5.1 we
 178 disprove this conjecture and prove in Theorem 14 that $L(4, n) = \Omega(n^{\Theta(\ln(n))})$. This shows
 179 that 4-deviations are responsible for a sudden complexity increase, as no polynomial
 180 bounds exist for $L(4, n)$. *This explains why in the foreword it takes an exponential time*
 181 *for the organizers of FUN to schedule the groups.*

2 Notations

183 **Conflict graph.** We refer to [2] for standard graph terminology. For the remaining of the
 184 paper, we suppose that we are given a *conflict graph* $G^- = (V, E)$ where V is the set of
 185 vertices (called users or players in the introduction) and edges represent the incompatibility
 186 relation (*i.e.*, an edge means that the two users are enemies). The number of vertices is
 187 denoted by $n = |V|$. The independence number of G^- , denoted $\alpha(G^-)$, is the maximum
 188 cardinality of an independent set in G^- . In particular, if $\alpha(G^-) = n$ then the conflict graph
 189 is edgeless and we denote it by $G^\emptyset = (V, E = \emptyset)$ and call it the empty graph.

190 **Partitions and utilities.** We consider any partition $P = X_1, \dots, X_i, \dots, X_n$ of the
 191 vertices into n independent sets X_i called groups (colors in coloring games), with some of
 192 them being possibly empty. In particular, two enemies are not in the same group. We rank
 193 the groups by non increasing size, that is $|X_i| \geq |X_{i+1}|$. For any $1 \leq i \leq n$ and for any

194 $v \in X_i$, the *utility* of vertex v is the number of other vertices in the same group as it, that
195 is $|X_i| - 1$.

196 We use in our proofs two alternative representations of the partition P . The *partition*
197 *vector* associated to P is defined as $\vec{\Lambda}(P) = (\lambda_n(P), \dots, \lambda_1(P))$, where $\lambda_i(P)$ is the number
198 of groups of size i . The *integer partition* associated to P is defined as $Q = (q_1, q_2, \dots, q_n)$
199 such that $q_1 \geq q_2 \geq \dots \geq q_n \geq 0$ and $\sum_{i=1}^n q_i = n$, where $q_i = |X_i|$.

200 In the example of Figure 1(a) we have a partition P of the 12 vertices into 4 groups
201 each of size 3 and so $\lambda_3(P) = 4$ and $\lambda_i(P) = 0$ for $i \neq 3$; in other words $\vec{\Lambda}(P) =$
202 $(0, 0, 0, 0, 0, 0, 0, 0, 0, 4, 0, 0)$. The integer partition $Q(P) = (3, 3, 3, 3, 0, 0, 0, 0, 0, 0, 0, 0)$. In
203 the example of Figure 1(b) we have a partition P' of the 12 vertices into one group of
204 size 4 and 4 groups each of size 2 and so $\lambda_4(P') = 1$, $\lambda_2(P') = 4$ and $\lambda_i(P') = 0$
205 for $i \neq 2, 4$; in other words $\vec{\Lambda}(P') = (0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 4, 0)$. The integer partition
206 $Q(P') = (4, 2, 2, 2, 2, 0, 0, 0, 0, 0, 0, 0)$.

207 **k-deviations and k-stability.** We can think of a k -deviation as a move of at most k
208 vertices which leave the groups to which they belong in P , to join another group (or create
209 a new group) with the necessary condition that each vertex strictly increases its utility,
210 thereby leading to a new partition P' . A k -stable partition is simply a partition for which
211 there exists no k -deviation. We write $L(k, G^-)$, resp. $L(k, n)$, for the length of a longest
212 sequence of k -deviations to reach a stable partition in G^- , resp. in any conflict graph with
213 n vertices. Recall that we start with the partition consisting of n groups of size 1, that is,
214 $\vec{\Lambda}(P) = (\dots, 0, 0, 0, n)$.

215 We next define a natural vector representation for k -deviations. The *difference vector* $\vec{\varphi}$
216 associated to a k -deviation φ from P to P' is equal to $\vec{\varphi} = \vec{\Lambda}(P') - \vec{\Lambda}(P)$. In concluding
217 this section, we define the difference vectors for some of the k -deviations used in our proofs:

- 218 ■ $\vec{\alpha}[p, q]$, the 1-deviation where a vertex leaves a group of size $q + 1$ for a group of size
219 $p - 1$ (valid when $p \geq q + 2$). In that case $\alpha_p = 1, \alpha_{p-1} = -1, \alpha_{q+1} = -1, \alpha_q = -1$, and
220 $\alpha_i = 0$ for any $i \notin \{q, q + 1, p - 1, p\}$ (we omit for ease of reading the brackets $[p, q]$).
- 221 ■ $\vec{\gamma}[p]$, the 3-deviation where one vertex in each of 3 groups of size $p - 1$ moves to a group
222 of size $p - 3$ to form a new group of size p (valid if there are at least 3 groups of size
223 $p - 1$ and one of size $p - 3$). In that case $\gamma_p = 1, \gamma_{p-1} = -3, \gamma_{p-2} = 3, \gamma_{p-3} = -1$, and
224 $\gamma_i = 0$ for any $i \notin \{p - 3, p - 2, p - 1, p\}$.
- 225 ■ $\vec{\delta}[p]$, the 4-deviation where one vertex in each of 4 groups of size $p - 1$ moves to a group
226 of size $p - 4$ to form a new group of size p (valid if there are at least 4 groups of size
227 $p - 1$ and one of size $p - 4$). In that case $\delta_p = 1, \delta_{p-1} = -4, \delta_{p-2} = 4, \delta_{p-4} = -1$, and
228 $\delta_i = 0$ for any $i \notin \{p - 4, p - 2, p - 1, p\}$. As an example, the move from the partition of
229 Figure 1(a) to the partition of Figure 1(b), is a 4-deviation with difference vector $\vec{\delta}[4]$.

230 3 Preliminary results

231 In [14], the authors prove that there always exists a k -stable partition, but that it is NP-hard
232 to compute one (this result was also proved independently in [7]). In contrast, it can be
233 computed a k -stable partition in polynomial time for every fixed $k \leq 3$, by using simple
234 *better-response dynamics* [18, 7, 14]. The latter results question the role of the value of k in
235 the complexity of computing stable partitions.

236 Formally, a better-response dynamics proceeds as follows. We start from the trivial
237 partition P_1 consisting of n groups with one vertex in each of them. In particular, the
238 partition vector $\vec{\Lambda}(P_1)$ is such that $\lambda_1(P_1) = n$ and, for all other $j \neq 1$, $\lambda_j(P_1) = 0$.

239 Provided there exists a k -deviation with respect to the current partition P_i , we pick any
 240 one of these k -deviations φ and in so doing we obtain a new partition P_{i+1} . If there is no
 241 k -deviation, the partition P_i is k -stable. An algorithmic presentation is given in Algorithm 1.

Dynamics of the system (Algorithm 1)

Input: a positive integer $k \geq 1$, and a conflict graph G^- .

Output: a k -stable partition for G^- .

- 1: Let P_1 be the partition composed of n singletons groups.
 - 2: Set $i = 1$.
 - 3: **while** there exists a k -deviation for P_i **do**
 - 4: Set $i = i + 1$.
 - 5: Choose one k -deviation and compute the partition P_i after this k -deviation.
 - 6: Return the partition P_i .
-

242 We now prove in Proposition 1 that better-response dynamics can be used for computing
 243 a k -stable partition *for every fixed* $k \geq 1$ (but not necessarily in polynomial time). It
 244 shows that for every fixed $k \geq 1$, the problem of computing a k -stable partition is in the
 245 complexity class PLS (Polynomial Local Search), that is conjectured to lie strictly between
 246 P and NP [13]. Recall that the problem becomes NP-hard when k is part of the input.

247 ► **Proposition 1.** For any $k \geq 1$, for any conflict graph G^- , Algorithm 1 converges to a
 248 k -stable partition.

249 **Proof.** Let P_i, P_{i+1} be two partitions for G^- such that P_{i+1} is obtained from P_i after some
 250 k -deviation φ . Let S be the set of vertices which move ($|S| \leq k$) and let j be the size of
 251 the group they join ($j = 0$ if they create a new group). Then, the new group obtained has
 252 size $p = j + |S|$. Note that all the vertices of S have increased their utilities and so, they
 253 belonged in P_i to groups of size $< p$. Therefore, the coordinates of the difference vector
 254 $\vec{\varphi}$ satisfy $\varphi_p = 1$ and $\varphi_j = 0$ for $j > p$, and so $\vec{\Lambda}(P_i) <_L \vec{\Lambda}(P_{i+1})$ where $<_L$ is the
 255 lexicographical ordering. Finally, as the number of possible partition vectors is finite, we
 256 obtain the convergence of Algorithm 1. ◀

257 An instrumental observation for our next proofs is the following:

258 ► **Observation 1.** $L(k, n)$ is always attained on the empty conflict graph G^\emptyset of order n .

259 Indeed, any sequence of k -deviations on a conflict graph G^- is also a sequence in the empty
 260 conflict graph with the same vertices. Note that the converse is not true as it can happen
 261 that some moves allowed in the empty conflict graph are not allowed in G^- as they bring
 262 two enemies in the same group.

263 Recall that we can associate to any partition $P = X_1, \dots, X_i, \dots, X_n$ of the vertices the
 264 integer partition $Q = (q_1, q_2, \dots, q_n)$ such that $q_1 \geq q_2 \geq \dots \geq q_n \geq 0$ and $\sum_{i=1}^n q_i = n$ by
 265 letting $q_i = |X_i|$. The converse is not true in general; as example it suffices to consider a
 266 partition with $q_1 > \alpha(G^-)$. However the converse is true when the conflict graph is empty;
 267 indeed it suffices to associate to an integer partition any partition of the vertices obtained
 268 by putting in the group X_i a set of q_i vertices.

269 We can now use the value of the number p_n of the number of integer partitions (see [11])
 270 to obtain the following proposition which follows from Proposition 1.

271 ► **Proposition 2.** Algorithm 1 reaches a stable partition in at most $p_n = \Theta((e^{\pi\sqrt{\frac{2n}{3}}})/n)$ steps.

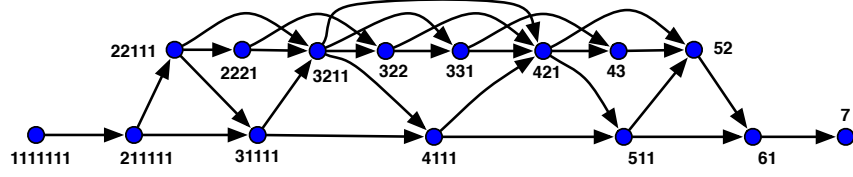


Figure 2 The lattice of integer partitions for $n = 7$.

Note that this is already far less than 2^n , which was shown to be the best upper bound that one can obtain for $k \geq 4$ when using an additive potential function [14].

4 Analysis for $k \leq 2$

In [14], the authors proved that for $k \leq 2$, Algorithm 1 converges to a stable partition in at most a quadratic time (by using a potential function). Indeed when performing a 1-deviation $\vec{\alpha}[p, q]$, a vertex moves from a group of size $q + 1$ to a group of size $p - 1$ (with $p \geq q + 2$); the utility of this vertex increases by $p - q - 1$, the utility of the q other vertices of the group of size $q + 1$ decreases by 1, while the utility of the vertices of the group of size $p - 1$ increases by 1 and so the global utility increases by $2p - 2q - 2 \geq 2$ as $p \geq q + 2$. Furthermore, in a k -stable partition, the utility of a vertex is at most $n - 1$ and the global utility is at most $n(n - 1)/2$ and so $L(k, n) = O(n^2)$.

In the next subsections we improve this result as we completely solve this case and give the exact (non-asymptotic) value of $L(k, n)$ when $k \leq 2$. The gist of the proof is to use a partial ordering that was introduced in [3], and is sometimes called the dominance ordering.

4.1 Exact analysis for $k = 1$ and empty conflict graph

In [3] the author has defined an ordering over the integer partitions, sometimes called the dominance ordering which creates a lattice of integer partitions. This ordering is a direct application of the theory of majorization to integer partitions [17].

► **Definition 3.** (dominance ordering) Given two integer partitions of $n \geq 1$, $Q = (q_1, q_2, \dots, q_n)$ and $Q' = (q'_1, q'_2, \dots, q'_n)$, we say that Q' dominates Q if $\sum_{j=1}^i q'_j \geq \sum_{j=1}^i q_j$, for all $1 \leq i \leq n$.

The example of Figure 2 shows the dominance lattice for $n = 7$. We did not write in the figure the integers equal to 0. Now we will show that there is a one to one mapping between chains in the dominance lattice and sequences of 1-deviations in the empty conflict graph.

The two next lemmas show that there is a one to one mapping between chains in the dominance lattice and sequences of 1-deviations in the empty conflict graph.

► **Lemma 4.** Let P be a partition of the vertices and P' be the partition obtained after a 1-deviation φ . Then, the integer partition $Q' = Q(P')$ dominates $Q = Q(P)$.

Proof. In the 1-deviation φ a vertex v moves from a group X_k to a group X_j with $q_j = |X_j| \geq q_k = |X_k|$. W.l.o.g. we can suppose that the groups (ranked in non increasing order of size) are ranked in a such a way that X_j is the first group with size $|X_j|$ and X_k the last group with size $|X_k|$. Thus, the integer partition $Q(P)$ associated to P satisfies $q_1 \geq q_2 \geq \dots \geq q_{j-1} > q_j \geq q_{j+1} \geq \dots \geq q_k > q_{k+1} \geq \dots \geq q_n$. After the move the groups of P' are the same as those of P except we have replaced X_j with the group $X_j \cup v$ and X_k with $X_k - v$. Therefore the integer partition Q' associated to P' has the same elements as Q

except $q'_j = q_j + 1$ and $q'_k = q_k - 1$ and so Q' dominates Q . Note that this lemma holds for any conflict graph. ◀

In the case $n = 7$, consider the partition P with one group of size 3, one of size 2 and two of size 1. The integer partition associated to P is $Q = (3, 2, 1, 1, 0, 0, 0)$. Let φ be the 1-deviation where a vertex in the group of size 1 moves to the group of size 3. We obtain the partition P' with one group of size 4, one of size 2 and one of size 1. The integer partition associated to P' is $Q' = (4, 2, 1, 0, 0, 0, 0)$ which dominates Q .

► **Lemma 5.** *Let G^0 be the empty conflict graph and let Q, Q' be two integer partitions of $n = |V|$ such that Q' dominates Q . For any partition P associated to Q , there exists another partition P' associated to Q' such that P' is obtained from P by doing a sequence of 1-deviations.*

Proof. As proved in [3], we have that if Q' dominates Q then there is a finite sequence of integer partitions $Q^0, \dots, Q^r, \dots, Q^s$, with $Q = Q^0$ and $Q' = Q^s$ such that for each $0 \leq r < s$, Q^{r+1} dominates Q^r and differs from it only in two elements j_r and k_r with $q_{j_r}^{r+1} = q_{j_r}^r + 1$ and $q_{k_r}^{r+1} = q_{k_r}^r - 1$.

The proof is now by induction on r , starting from any partition $P^0 = P$ associated to Q . For $r > 0$, we consider the partition P^r associated to Q^r . Recall that Q^r and Q^{r+1} differ only in the two groups X_{j_r} and X_{k_r} . As $q_{j_r}^{r+1} = q_{j_r}^r + 1$ and $q_{k_r}^{r+1} = q_{k_r}^r - 1$, P^{r+1} can be obtained from P^r by moving a vertex from X_{k_r} to X_{j_r} . This move is valid as the conflict graph is empty. (Note that the lemma is not valid for a general conflict graph.) ◀

As an example, consider the two integer partitions $Q = (2, 2, 2, 1, 0, 0, 0)$ and $Q' = (5, 1, 1, 0, 0, 0, 0)$ where Q' dominates Q . The sequence of integer partitions is $Q_0 = Q$, $Q_1 = (3, 2, 1, 1, 0, 0, 0)$, $Q_2 = (4, 1, 1, 1, 0, 0, 0)$, $Q_3 = (5, 1, 1, 0, 0, 0, 0)$. Partition P^1 is obtained from P^0 by moving a vertex of a group of size 2 to another group of size 2. Then, P^2 is obtained by moving a vertex of the group of size 2 to the group of size 3 and P' is obtained from P^2 by moving a vertex of one group of size 1 to that of size 4.

In summary we conclude that a sequence of 1-deviations with an empty conflict graph corresponds to a chain of integer partitions, and vice versa. Therefore, by Observation 1, the length of the longest sequence of 1-deviations with an empty conflict graph is the same as the length of the longest chain in the dominance lattice of integer partitions. Since it has been proven in [9] that for $n = \frac{m(m+1)}{2} + r$, the longest chain in the Dominance Lattice has length $2\binom{m+1}{3} + mr$, we obtain the exact value for $L(1, n)$.

► **Theorem 6.** *Let m and r be the unique non negative integers such that $n = \frac{m(m+1)}{2} + r$, and $0 \leq r \leq m$. Then, $L(1, n) = 2\binom{m+1}{3} + mr$.*

We note that the proof in [9] is not straightforward. One can think that the longest chain is obtained by taking among the possible 1-deviations the one which leads to the smallest partition in the lexicographic order. Unfortunately this is not true. Indeed let $n = 9$. After 6 steps we get the integer partition $(3, 3, 2, 1, 0, 0, 0, 0, 0)$. Then, by choosing the 1-deviation that gives the smallest partition (in the lexicographic order), we get the partition $(3, 3, 3, 0, 0, 0, 0, 0, 0)$ and then $(4, 3, 2, 0, 0, 0, 0, 0, 0)$. But there is a longer chain of length 3 from $(3, 3, 2, 1, 0, 0, 0, 0, 0)$ to $(4, 3, 2, 0, 0, 0, 0, 0, 0)$ namely $(4, 2, 2, 1, 0, 0, 0, 0, 0)$, $(4, 3, 1, 1, 0, 0, 0, 0, 0)$, $(4, 3, 2, 0, 0, 0, 0, 0, 0)$. However the proof in [9] implies that the following simple construction works for any n .

► **Proposition 7.** A longest sequence of 1-deviations in the empty conflict graph is obtained by choosing, at a given step, among all the possible 1-deviations, any one of which leads to the smallest increase of the global utility.

Proof. We need to introduce the terminology of [9] (this will not be used elsewhere in the paper). Note that they consider their chains starting from the end, and so, we need to reverse the steps in their construction in order to make them correspond with 1-deviations.

- A V -step is corresponding to a user leaving her group of size p for another group of size p , thereby increasing her utility from $p - 1$ to p ; in other words, the deviation vector of such 1-deviation is $\vec{\alpha}[p + 1, p - 1]$ for some p .
- An H -step is corresponding to a user leaving her group X_i of size p for another group X_j of size $p \leq |X_j| \leq p + 1$, but *only* if there is no other group of size p than X_i and (possibly) X_j ; in particular, if groups are ordered by decreasing size, this means that $j = i - 1$. The deviation vector of an H -step is either $\vec{\alpha}[p + 1, p - 1]$ (and then, X_i, X_j are the only two groups of size p) or $\vec{\alpha}[p + 2, p - 1]$ (and then X_i is the unique group of size p). Furthermore, note that an H -step can also be a V -step.

The relationship between V -steps, H -steps and our construction is as follows. At every 1-deviation, the global utility has to increase by at least two, and this is attained if and only if the deviation vector is $\vec{\alpha}[p + 1, p - 1]$ for some p ; equivalently, this move is corresponding to a V -step. Furthermore, if no such a move is possible, then the global utility has to increase by at least four, and this is attained if and only if the deviation vector is $\vec{\alpha}[p + 2, p - 1]$ for some p ; in such a case, there cannot exist any other group of size p than the one involved in the 1-deviation (otherwise, a V -step would have been possible), hence this move is corresponding to a H -step. As proved, *e.g.*, by Brylawski [3], starting from any integer partition with at least two summands (*i.e.*, $q_n \neq 1$), it is always possible to perform one of the two types of move defined above (these two types of move actually correspond to the two cases when an integer partition can cover another one). Therefore, our strategy leads to a sequence of 1-deviations where all the moves correspond to either a V -step or (only if the first type of move is not possible) to an H -step. By a commutativity argument (Lemma 3 in [9]) it can be proved that as soon as no move $\vec{\alpha}[p + 1, p - 1]$ (corresponding to V -steps) is possible for any p , every ulterior move of this type will correspond to both a V -step and an H -step simultaneously (*i.e.*, the deviation vector will be $\vec{\alpha}[p + 1, p - 1]$ for some p , and there will be no other groups of size p than the two groups involved in the 1-deviation). Therefore, the sequences we obtain are corresponding to a particular case of the so-called HV -chains in [9]. Finally, the main result in [9] is that every HV -chain is of maximum length. ◀

In what follows, we will reuse part of the construction in [9] for proving Theorem 10.

4.2 Analysis for $k = 2$

Interestingly we will prove that any 2-deviation can be replaced either by one or two 1-deviations and so we will prove in Theorem 9 that $L(2, n) = L(1, n)$.

► **Claim 8.** If the conflict graph G^- is empty, then any 2-deviation can be replaced either by one or two 1-deviations

Proof. Consider a 2-deviation which is not a 1-deviation. In that case case two vertices u_i and u_j leave their respective group X_i and X_j (which can be the same) to join a group X_k . Let $|X_i| \geq |X_j|$; in order for the utility of the vertices to increase, we should have $|X_k| \geq |X_i| - 1$ ($\geq |X_j| - 1$).

- Case 1: $|X_k| \geq |X_j|$. In that case the 2-deviation can be replaced by a sequence of two 1-deviations where firstly a vertex u_j leaves X_j to join X_k and then a vertex u_i leaves X_i to join the group $X_k \cup u_j$ whose size is now at least that of X_i .

396 ■ Case 2: $|X_k| = |X_i| - 1 = |X_j| - 1 = p - 2$ and $X_i = X_j$. In that case the effect of the
 397 2-deviation is to replace the group X_i of size $p - 1$ with a group of size $p - 3$ and to replace
 398 the group X_k of size $p - 2$ with a group of size p . Said otherwise, the difference vector
 399 $\vec{\varphi}$ associated to the 2-deviation has as non null coordinates $\varphi_p = 1, \varphi_{p-1} = -1, \varphi_{p-2} =$
 400 $-1, \varphi_{p-3} = 1$. We obtain the same effect by doing the 1-deviation $\vec{\alpha}[p-1, p-2]$ where
 401 a vertex leaves X_k to join X_i .
 402 ■ Case 3: $|X_k| = |X_i| - 1 = |X_j| - 1 = p - 2$ and $X_i \neq X_j$. In that case the effect of the
 403 2-deviation is to replace the 2 groups X_i and X_j of size $p - 1$ with two groups of size
 404 $p - 2$ and to replace the group X_k of size $p - 2$ with a group of size p . Said otherwise,
 405 the difference vector $\vec{\varphi}$ associated to the 2-deviation has as non null coordinates $\varphi_p =$
 406 $1, \varphi_{p-1} = -2, \varphi_{p-2} = 1$. We obtain the same effect by doing the 1-deviation $\vec{\alpha}[p-1, p-1]$
 407 where a vertex leaves X_j to join X_i .
 408 Note that the fact G^- is empty is needed for the proof. Indeed, in the case 2, it might happen
 409 that all the vertices of X_k have some enemy in X_i and so the 1-deviation we describe is not
 410 valid. Similarly, in case 3, it might happen that all the vertices of X_i have some enemy in
 411 X_j and so the 1-deviation we describe is not valid. \diamond

412 ► **Theorem 9.** $L(2, n) = L(1, n)$.

413 **Proof.** Clearly, $L(2, n) \geq L(1, n)$ as any 1-deviation is also a 2-deviation. By Observation 1,
 414 the value of $L(2, n)$ is obtained when the conflict graph G^- is empty. In that case, Claim 8
 415 implies that $L(2, n) \leq L(1, n)$. \blacktriangleleft

416 4.3 Analysis for $k \leq 2$ and a general conflict graph

417 Using the potential function introduced at the beginning of this section, Panagopoulou and
 418 Spirakis ([18]) proved that for every conflict graph G^- with independence number $\alpha(G^-)$,
 419 the convergence of the dynamics is in $\mathcal{O}(n\alpha(G^-))$. Indeed as we have seen each 1-deviation
 420 increases the global utility by at least 2. But the global utility of a stable partition is at most
 421 $n(\alpha(G^-) - 1)$ as the groups have maximum size $\alpha(G^-)$. If the conflict graph is empty we
 422 have seen that $L(1, n) = \Theta(n^{3/2})$ that is in that case $\mathcal{O}(n\sqrt{\alpha(G^-)})$. This leads one of us ([6],
 423 page 131) to conjecture that in the case of 1-deviations the worst time of convergence of the
 424 dynamics is $\mathcal{O}(n\sqrt{\alpha(G^-)})$. We disprove the conjecture by proving the following theorem:

425 ► **Theorem 10.** For $n = \binom{m+1}{2}$, there exists a conflict graph G^- with $\alpha(G^-) = m = \Theta(\sqrt{n})$
 426 and a sequence of $\binom{m+1}{3}$ valid 1-deviations, that is a sequence of $\Omega(n^{\frac{3}{2}}) = \Omega(n\alpha(G^-))$
 427 1-deviations.

428 **Proof.** We will use part of the construction of Greene and Kleitman ([9]). Namely they
 429 prove that, if $n = \binom{m+1}{2}$, there is a sequence of $\binom{m+1}{3}$ 1-deviations transforming the partition
 430 P_1 consisting of n groups each of size 1 (the coordinates of $\vec{\lambda}(P_1)$ satisfy $\lambda_1 = n$) into the
 431 partition P_m consisting of m groups, one of each possible size i for $1 \leq i \leq m$ (the coordinates
 432 of $\vec{\lambda}(P_m)$ satisfy $\lambda_i = 1$ for $1 \leq i \leq m$). Furthermore they prove that the moves used are
 433 V -steps (see the proof of proposition 7 that is $\vec{\alpha}[p+1, p-1]$ for some p (one vertex leaves
 434 a group of size p to join a group of the same size p). One can note that in such a move the
 435 utility increases only by 2 and as the total utility of P_m is $\sum_{i=1}^m i(i-1) = (m+1)m(m-1)/3$
 436 the number of moves is $(m+1)m(m-1)/6$.

437 The conflict graph of the counterexample will consist of m complete graphs $K^j, 1 \leq j \leq$
 438 m where K^j has exactly j vertices. An independent set is therefore formed by taking at
 439 most one vertex in each K^j and $\alpha(G^-) = m$. We will denote the elements of K^j by $\{x_i^j\}$

with $1 \leq i \leq j \leq m$. The group of P_m of size i will be $X_i = \bigcup x_i^j$ with $m+1-i \leq j \leq m$. So these groups are independent sets.

Recall that $n = m(m+1)/2$. For each $p, 1 \leq p \leq m$ let us denote by P_p the partition consisting of 1 group of each size i for $1 \leq i \leq p$ and $n - p(p+1)/2$ groups of size 1 (said otherwise the coordinates of $\vec{\lambda}(P_p)$ satisfy $\lambda_i = 1$ for $2 \leq i \leq p$ and $\lambda_1 = 1 + n - p(p+1)/2$). We will now describe the sequence $\vec{\sigma}[p-1]$ of $p(p-1)/2$ 1-deviations which transform the partition P_{p-1} into P_p . One way to do the Greene-Keitman sequence is obtained by doing successively the sequences $\sum_{p=2}^m \vec{\sigma}[p-1]$. More precisely we will prove by induction the following fact:

► **Claim 11.** There exists a sequence $\vec{\sigma}[p-1]$ of $p(p-1)/2$ valid 1-deviations which transform the partition P_{p-1} into P_p such that after this sequence the group $X_i[p]$ of size i , $1 \leq i \leq p$ contains exactly the vertices $X_i[p] = \bigcup x_{i+m-p}^j$ with $m+1-i \leq j \leq m$.

Proof. (the reader can follow the construction on the example after)

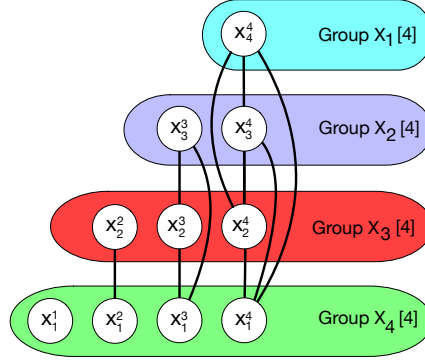
We suppose we have built the sequence till $p-1$ and that, for $1 \leq i \leq p-1$, $X_i[p-1] = \bigcup x_{i+m-p+1}^j$ with $m+1-i \leq j \leq m$. In a first phase we consider the subpartition of $n-p+1$ elements obtained by removing the group $X_{p-1}[p-1]$. Namely, this above subpartition consists of the groups $X_i[p-1]$ for $1 \leq i \leq p-2$ and groups of size 1. In particular, the subpartition is isomorphic to P_{p-2} with $p-1$ singleton groups removed. Our construction ensure that these $p-1$ singleton groups that are missing are not used for $\vec{\sigma}[p-2]$. So, we can do the transformation $\vec{\sigma}[p-2]$ consisting of $(p-1)(p-2)/2$ valid moves on the partition of $n-p+1$ elements not contained in $X_{p-1}[p-1]$. It gives rise to the groups $X_i[p] = X_{i-1}[p-1] + x_{i+m-p}^{m+1-i}$. Note that at this stage we have two groups of size $p-1$, namely the original one $X_{p-1}[p-1]$ and the new one constructed $X_{p-1}[p]$. The second phase consists in doing $p-1$ successive 1-deviations with the vertex x_m^{m+1-p} . More precisely we move this vertex to the group $X_1[p]$ created in the first phase, then from this group to $X_2[p]$ and so on till $X_{p-2}[p]$ and finally from $X_{p-2}[p]$ to the original $X_{p-1}[p-1]$. The moves are valid as we move a vertex from K^{m+1-p} and the groups did not contain any vertex of this complete graph. Groups created in the first phase are eventually left unchanged as x_m^{m+1-p} joins this group and then leaves it. Finally we have constructed a new group $X_p[p] = X_{p-1}[p-1] \cup x_m^{m+1-p}$. The groups are exactly those described in the claim. ◊

To end the proof of Theorem 10, it suffices to note that the groups X_i form an independent set and that after $\sum_{p=2}^m \vec{\sigma}[p-1]$ we have obtained the desired groups of P_m which gives the counterexample. ◀

Example for $m = 4$. (See Figure 3.)

- After $\vec{\sigma}[1]$, we have the 2 groups $X_2[2] = x_4^4 \cup x_4^3$ and $X_1[2] = x_3^4$.
- First phase of $\vec{\sigma}[2]$: we do the move of $\vec{\sigma}[1]$ on the vertices not in $X_2[2]$ and create the groups $X_2[3] = x_3^4 \cup x_3^3$ and $X_1[3] = x_2^4$.
- Second phase of $\vec{\sigma}[2]$: now we move x_4^2 to $X_1[3]$ and then from $X_1[3]$ to the original $X_2[2] = x_4^4 \cup x_4^3$, thereby creating the group $X_3[3] = x_4^4 \cup x_4^3 \cup x_4^2$.
- First phase of $\vec{\sigma}[3]$: we do the 3 moves of $\vec{\sigma}[2]$ on the vertices not in $X_3[3]$ and create the groups $X_3[4] = x_3^4 \cup x_3^3 \cup x_3^2$, $X_2[4] = x_2^4 \cup x_2^3$, $X_1[4] = x_1^4$.
- Second phase of $\vec{\sigma}[3]$: now we move x_4^1 to $X_1[4]$, then from $X_1[4]$ to $X_2[4]$ and finally from $X_2[4]$ to the original $X_3[3] = x_4^4 \cup x_4^3 \cup x_4^2$, thereby creating the group $X_4[4] = x_4^4 \cup x_4^3 \cup x_4^2 \cup x_4^1$.

We can prove a theorem analogous to Theorem 10 for any independence number $\alpha(G^-)$.



■ **Figure 3** Illustration for Example 4.

486 ► **Theorem 12.** For any $\alpha = \mathcal{O}(\sqrt{n})$, there exists a conflict graph G^- with n vertices and
 487 independence number $\alpha(G^-) = \alpha$, and a sequence of at least $\Omega(n\alpha)$ 1-deviations to reach a
 488 stable partition.

489 **Proof.** Let G_0^- be the graph of Theorem 10 for $m = \alpha$. G_0^- has $n_0 = \mathcal{O}(\alpha^2)$ vertices,
 490 independence number α , and furthermore there exists a sequence of $\Theta(\alpha^3)$ valid 1-deviations
 491 for G_0^- . Let G^- be the graph obtained by taking the complete join of $k = n/n_0$ copies of
 492 G_0^- (i.e., we add all possible edges between every two copies of G_0^-). By construction, G^-
 493 has order $n = kn_0 = \mathcal{O}(n\alpha^2)$ and the same independence number α as G_0^- . Furthermore,
 494 there exists a sequence of $k\Theta(\alpha^3) = \Omega(n\alpha)$ valid 1-deviations for G^- . ◀

495 Note that in any 2-deviation the global utility increases by at least 2 and so the number
 496 of 2 deviations when the conflict graph has independence number $\alpha(G^-)$ is also at most
 497 $\mathcal{O}(n\alpha(G^-))$. This bound is attained by using only 1-deviations as proved in Theorem 12,
 498 which is also valid for $k = 2$.

499 5 Lower bounds for $k > 2$

500 The classical dominance ordering does not suffice to describe all k -deviations as soon as
 501 $k \geq 3$. As noted before, there is only one k -stable partition P_{max} in the empty conflict
 502 graph G^\emptyset , namely the one consisting of one group of size n , with integer partition $Q_{max} =$
 503 $(n, 0, \dots, 0)$ and partition vector $(1, 0, \dots, 0)$. Let $d(Q)$ be the length of a longest sequence in
 504 the dominance lattice from the integer partition Q to the integer partition Q_{max} . For $k = 4$
 505 let P be the partition consisting of 4 groups of size 4 and one group of size 1 with integer
 506 partition $Q = (4, 4, 4, 4, 1)$. Apply the 4-deviation where one vertex of each group of size 4
 507 joins the group of size 1; it leads to the partition P' with integer partition $Q' = (5, 3, 3, 3, 3)$.
 508 Q is covered in the dominance lattice by the integer partition $(5, 4, 4, 3, 1)$ while Q' is at
 509 distance 3 from it via $(5, 4, 3, 3, 2)$ and $(5, 4, 4, 2, 2)$ and so $d(Q') = d(Q) + 2$.

510 Prior to our work, it was known $L(3, n) = \mathcal{O}(n^3)$ ([14]). But nothing proved that
 511 $L(3, n) > L(2, n)$, and in fact it was conjectured in [7] that both values are equal. Theorem 13
 512 proves for the first time that deviations can delay convergence and that the gap between
 513 $k = 2$ and $k = 3$ obtained from potential function is indeed justified. It was also conjectured
 514 in [14] that $L(k, n)$ was polynomial in n for k fixed. We disprove this conjecture and prove
 515 in Theorem 14 a much more significant result: 4-deviations are responsible for a sudden
 516 complexity increase, as we prove that no polynomial bounds exist for $L(4, n)$.

517 ► **Theorem 13.** $L(3, n) = \Omega(n^2)$.

518 ► **Theorem 14.** $L(4, n) = \Omega(n^{\Theta(\ln(n))})$.

519 The main idea of the proofs consists in doing repeated shifted sequences (called cascades)
 520 of deviations similar to the ones given in the example above. The proof of Theorems 13
 521 can be found in the appendix (Section 5.2). In the next section, we give the the proof of
 522 Theorem 14 for $k = 4$. We use sequences (cascades) of 4-deviations, called $\delta[p]$, and various
 523 additional tricks such that the repetition of the process by using cascades of cascades. Our
 524 motivation for using $\delta[p]$ as a basic building block for our construction is that it is the only
 525 type of 4-deviation which decreases the global utility.

526 5.1 Case $k = 4$. Proof of Theorem 14

527 **Definition of $\delta[p]$:** Consider a partition P containing at least 4 groups of size $p - 1$ and 1
 528 group of size $p - 4$. In the 4-deviation $\delta[p]$ one vertex in each of the 4 groups of size $p - 1$
 529 moves to the group of size $p - 4$ to form a new group of size p . The example given at the
 530 beginning of this section corresponds to the case $p = 5$. The coordinates of the associated
 531 difference vector (where we omit the bracket $[p]$ for ease of reading) are:

532

...	δ_p	δ_{p-1}	δ_{p-2}	δ_{p-3}	δ_{p-4}	...
...0	1	-4	4	0	-1	0...

533 Figure 4 gives a visual description of these cascades. Here we start with a sequence of
 534 t 4-deviations $\delta[p]$ represented by black rectangles ($t = 16$ in the figure). The cascade so
 535 obtained, called $\vec{\delta}^1[p, t]$, is represented in red. Then we do $(t - 2)$ such cascades represented
 536 by red rectangles getting the cascade $\vec{\delta}^2[p, t - 2]$ represented in yellow which contains
 537 $224 (= 16 \cdot 14)$ 4-deviations. We apply some 1-deviations to get a deviation called $\vec{\tau}^2[p]$
 538 with the so-called Nice Property enabling us to do recursive constructions. We do a cascade
 539 of these $\vec{\tau}^2[p]$ (shifted by 2) represented by yellow rectangles getting the blue cascade called
 540 $\vec{\tau}^3[p]$. We do a cascade of these $\vec{\tau}^3[p]$ (shifted by 3) represented by blue rectangles getting
 541 the green cascade called $\vec{\tau}^4[p]$ and we finally do a cascade of these $\vec{\tau}^4[p]$ (shifted by 5)
 542 represented by green rectangles getting the grey cascade called $\vec{\tau}^5[p]$. The reader has to
 543 realize that, in this example, $\vec{\tau}^5[p]$ contains 3 cascades $\vec{\tau}^4[p]$ each containing 5 cascades
 544 $\vec{\tau}^3[p]$ each consisting of 7 cascades $\vec{\tau}^2[p]$. Altogether the cascade $\vec{\tau}^5[p]$ of this example
 545 contains 23520 4-deviations $\delta[p]$.

546 **The cascade $\vec{\delta}^1[p, t]$:** we first do a cascade consisting of a sequence of t shifted 4-
 547 deviations $\delta[p], \delta[p - 1], \dots, \delta[p - t + 1]$, for some parameter t which will be chosen later to
 548 give the maximum number of 4-deviations.

549 The reader can follow the construction in Table 2 with $t = 7$. The coordinates $\vec{\delta}^1[p, t]$,
 550 are given in Claim 15 and Table 10. We note that there are lot of cancellations and only 8
 551 non zero coordinates. Indeed consider the groups of size $p - i$ for $4 \leq i \leq t - 1$; we have
 552 deleted such a group when doing the 4-deviation $\vec{\delta}[p + 4 - i]$, then created 4 such groups
 553 with $\vec{\delta}[p + 2 - i]$, then deleted 4 such groups with $\vec{\delta}[p + 1 - i]$, and finally created one with
 554 $\vec{\delta}[p - i]$. The reader can follow these cancellations in Table 2 for $i = 4, 5, 6$. The variation
 555 of the number of groups of a given size $p - i$ (which correspond to the coordinate δ_{p-i}^1) is
 556 obtained by summing the coefficients appearing in the corresponding column and so is 0 for
 557 $p - 4, p - 5, p - 6$.

558 ► **Claim 15.** For $3 \leq t \leq p - 3$, the coordinates of the cascade $\vec{\delta}^1[p, t] = \sum_{i=0}^{t-1} \vec{\delta}[p - i]$
 559 satisfy: $\delta_p^1 = 1$, $\delta_{p-1}^1 = -3$, $\delta_{p-2}^1 = 1$, $\delta_{p-3}^1 = 1$, $\delta_{p-t}^1 = -1$, $\delta_{p-t-1}^1 = 3$, $\delta_{p-t-2}^1 = -1$,
 560 $\delta_{p-t-3}^1 = -1$, and $\delta_j^1 = 0$ for all the others j (see Table 10).



■ **Figure 4** Cascades of cascades.

561 *Proof.* We have $\delta_j^1 = \sum_{i=0}^{t-1} \delta_j[p-i]$. For a given j , $\delta_j[p-i] = 0$ except for the following values
 562 of i such that $0 \leq i \leq t-1$: $i = p-j$ where $\delta_j[j] = 1$; $i = p-j-1$ where $\delta_j[j+1] = -4$;
 563 $i = p-j-2$ where $\delta_j[j+2] = 4$; $i = p-j-4$ where $\delta_j[j+4] = -1$ (in the table it
 564 corresponds to the non zero values in a column, whose number is at most 4). Therefore, for
 565 $j > p$: $\delta_j^1 = 0$; $\delta_p^1 = 1$; $\delta_{p-1}^1 = -4 + 1 = -3$; $\delta_{p-2}^1 = 4 - 4 + 1 = 1$; $\delta_{p-3}^1 = 0 + 4 - 4 + 1 = 1$;
 566 for $p-4 \geq j \geq p-t+1$, $\delta_{p-j}^1 = -1 + 0 + 4 - 4 + 1 = 0$; $\delta_{p-t}^1 = -1 + 0 + 4 - 4 = -1$;
 567 $\delta_{p-t-1}^1 = -1 + 0 + 4 = 3$; $\delta_{p-t-2}^1 = -1 + 0 = -1$; $\delta_{p-t-3}^1 = -1$ and, for $j \leq p-t-4$, $\delta_j^1 = 0$.
 568 \diamond

569 **Validity of the cascades.** We have to see when the cascades are valid, that is, to
 570 determine how many groups we need at the beginning. For the cascade $\vec{\delta}^1[p, t]$ we note
 571 that the coordinates of any subsequence of the cascade, *i.e.*, the coordinates of some $\vec{\delta}^1[p, r]$,
 572 are all at least -1 except δ_{p-1}^1 : which is -4 when $r = 1$ and then -3 . Therefore such a
 573 cascade is valid as soon as we have at least 4 groups of size $p-1$ and one group of each
 574 other size $p-i$ ($2 \leq i \leq t+3$). To deal in general with the validity of cascades let us now
 575 introduce the notion of *h-balanced sequence*.

576 ► **Definition 16.** Let h be a positive integer and let $\vec{\Phi} = \sum_{j=1}^s \vec{\varphi}^j$ be a cascade consisting
 577 of s k -deviations. We call this cascade *h-balanced* if, for any $1 \leq i \leq s$, the sum of the i

XX:16 How long does it take for all users in a social network to choose their communities?

	...0	p	p-1	p-2	p-3	p-4	p-5	p-6	p-7	p-8	p-9	p-10	0...
$\delta[p]$...0	1	-4	4	0	-1	0	0	0	0	0	0	0...
$+\delta[p-1]$...0	0	1	-4	4	0	-1	0	0	0	0	0	0...
$+\delta[p-2]$...0	0	0	1	-4	4	0	-1	0	0	0	0	0...
$+\delta[p-3]$...0	0	0	0	1	-4	4	0	-1	0	0	0	0...
$+\delta[p-4]$...0	0	0	0	0	1	-4	4	0	-1	0	0	0...
$+\delta[p-5]$...0	0	0	0	0	0	1	-4	4	0	-1	0	0...
$+\delta[p-6]$...0	0	0	0	0	0	0	1	-4	4	0	-1	0...
$= \vec{\delta}^1[p, 7]$...0	1	-3	1	1	0	0	0	-1	3	-1	-1	0...

■ **Table 2** Computation of $\delta^1[p, 7]$.

...	δ_p^1	δ_{p-1}^1	δ_{p-2}^1	δ_{p-3}^1	...	δ_{p-t}^1	δ_{p-t-1}^1	δ_{p-t-2}^1	δ_{p-t-3}^1	...
...0	1	-3	1	1	0...0	-1	3	-1	-1	0...

■ **Table 3** Difference vector $\delta^1[p, t]$.

578 first vectors, namely $\sum_{j=1}^i \vec{\varphi}^j$, has all its coordinates greater than or equal to $-h$.

579 For example, the cascade $\vec{\delta}^1[p, t]$ described before is 4-balanced. The interest of this
 580 notion lies in the following fact: Let p_{\max} be the largest index j that satisfies $\vec{\Phi}_j \neq 0$. Then,
 581 if we start from a partition with at least h groups of each size j , for $1 \leq j \leq p_{\max}$, an
 582 h -balanced sequence is valid.

583 Note that a sequence is itself composed of sub-sequences and the following lemma will
 584 be useful to bound the value h of a sequence.

585 ► **Lemma 17.** *Let $\vec{\Phi}^1$ be an h_1 -balanced sequence and $\vec{\Phi}^2$ be an h_2 -balanced sequence. Then,*
 586 *$\vec{\Phi}^1 + \vec{\Phi}^2$ is a $(\max\{h_1, h_2 - \min_i \Phi_i^1\})$ -balanced sequence.*

587 *Proof.* As $\vec{\Phi}^1$ is h_1 -balanced, the coordinates of any subsequence of $\vec{\Phi}^1$ are greater than or
 588 equal to $-h_1$. Consider a subsequence $\vec{\Phi}^1 + \vec{\Phi}^3$ where $\vec{\Phi}^3$ is a subsequence of $\vec{\Phi}^2$. The j -th
 589 coordinate is $\Phi_j^1 + \Phi_j^3$; by definition $\Phi_j^3 \geq -h_2$ and so $\Phi_j^1 + \Phi_j^3 \geq \Phi_j^1 - h_2 \geq \min_i \Phi_i^1 - h_2$. \diamond

590 **The cascade $\vec{\delta}^2[p, t-2]$:** We do now the following sequence of $t-2$ cascades $\vec{\delta}^2[p, t-2]$
 591 $= \sum_{i=0}^{t-3} \vec{\delta}^1[p-i, t]$. Altogether we have a sequence of $t(t-2)$ 4-deviations. There are
 592 a lot of cancellations and in fact, as shown in Claim 18, $\vec{\delta}^2[p, t-2]$ has only 10 non zero
 593 coordinates. Table 4 describes an example of computation of $\vec{\delta}^2[p, t-2]$ with $t = 7$.

	...	p	p-1				p-5			p-9				p-13	p-14	...
$\delta^1[p, 7]$...0	1	-3	1	1	0	0	0	-1	3	-1	-1	0	...		
$+\delta^1[p-1, 7]$...0	0	1	-3	1	1	0	0	0	-1	3	-1	-1	0	...	
$+\delta^1[p-2, 7]$...0	0	0	1	-3	1	1	0	0	0	-1	3	-1	-1	0	...
$+\delta^1[p-3, 7]$...0	0	0	0	1	-3	1	1	0	0	0	-1	3	-1	-1	0
$+\delta^1[p-4, 7]$...0	0	0	0	0	1	-3	1	1	0	0	0	-1	3	-1	-1
$= \vec{\delta}^2[p, 5]$...0	1	-2	-1	0	0	-1	2	0	2	1	0	0	1	-2	-1

■ **Table 4** Computation of $\delta^2[p, 5]$.

► **Claim 18.** For $3 \leq t \leq \frac{p}{2}$, the coordinates of the cascade $\vec{\delta}^2[p, t-2] = \sum_{i=0}^{t-3} \vec{\delta}^1[p-i, t]$ satisfy: $\delta_p^2 = 1$, $\delta_{p-1}^2 = -2$, $\delta_{p-2}^2 = -1$, $\delta_{p-t+2}^2 = -1$, $\delta_{p-t+1}^2 = 2$, $\delta_{p-t}^2 = 2$, $\delta_{p-t-2}^2 = 1$, $\delta_{p-2t+2}^2 = 1$, $\delta_{p-2t+1}^2 = -2$, $\delta_{p-2t}^2 = -1$, and $\delta_j^2 = 0$ for all the others j (see Table 5). Furthermore this cascade is 4-balanced.

...	δ_p^2	δ_{p-1}^2	δ_{p-2}^2	...	δ_{p-t+2}^2	δ_{p-t+1}^2	δ_{p-t}^2	δ_{p-t-1}^2	δ_{p-t-2}^2	...	δ_{p-2t+2}^2	δ_{p-2t+1}^2	δ_{p-2t}^2	...
0	1	-2	-1	0	-1	2	0	2	1	0	1	-2	-1	0

■ **Table 5** Difference vector $\delta^2[p, t-2]$.

Proof. We have $\delta_j^2 = \sum_{i=0}^{t-3} \delta_j^1[p-i, t]$. Using the values of $\delta_j^1[p-i, t]$ given in Claim 23, we get that: for $j > p$, $\delta_j^2 = 0$; $\delta_p^2 = 1$; $\delta_{p-1}^2 = -3 + 1 = -2$; $\delta_{p-2}^2 = 1 - 3 + 1 = -1$; for $p-3 \geq j \geq p-t+3$, $\delta_j^2 = 1 + 1 - 3 + 1 = 0$; $\delta_{p-t+2}^2 = 1 + 1 - 3 = -1$; $\delta_{p-t+1}^2 = 1 + 1 = 2$; $\delta_{p-t}^2 = -1 + 1 = 0$; $\delta_{p-t-1}^2 = 3 - 1 = 2$; $\delta_{p-t-2}^2 = -1 + 3 - 1 = 1$; for $p-t-3 \geq j \geq p-2t+3$, $\delta_j^2 = -1 - 1 + 3 - 1 = 0$; $\delta_{p-2t+2}^2 = -1 - 1 + 3 = 1$; $\delta_{p-2t+1}^2 = -1 - 1 = -2$, $\delta_{p-2t}^2 = -1$, and for $j < p-2t$, $\delta_j^2 = 0$.

Using Lemma 17 we get that $\vec{\delta}^2[p, t-2]$ is 7-balanced; but a careful analysis shows that this sequence is in fact 4-balanced. Indeed we will prove by induction that $\vec{\delta}^2[p, r] = \sum_{i=0}^{r-1} \vec{\delta}^1[p-i, t]$ is 4-balanced for any $r \leq t-3$. That is true for $r = 1$, as $\vec{\delta}^1[p, t]$ is 4-balanced. Suppose that it is true for r . We have $\vec{\delta}^2[p, r+1] = \vec{\delta}^2[p, r] + \vec{\delta}^1[p-r-1, t]$. All the coordinates of $\vec{\delta}^2[p, r]$ are by the computation above at least -3 , and the coordinates of $\vec{\delta}^1[p-r-1, t]$ are greater than -1 except for $j = p-r-2$ where $\delta_{p-r-2}^1[p-r-1] = -4$; but $\delta_{p-r-2}^2[p, r] = 1$ (case $r = 1$) or 2 (case $r > 1$) and so all the coordinates of $\vec{\delta}^2[p, r+1]$ are at least -4 . \diamond

At this stage we could continue and do a cascade of $\vec{\delta}^2[p, t-2]$ but there is no more the phenomenon of cancellation. In fact we will use the following “symmetrization” trick. We will transform the cascade $\vec{\delta}^2[p, t-2]$ into a sequence $\vec{\zeta}^2[p]$ by doing some sequence of 1-deviations whose coordinates are given in Claim 19. The sequence obtained has only 8 non zero coefficients (4 with values 1 and 4 with values -1) arranged in a very symmetric nice way (that we will call *Nice Property*). Furthermore we will be able to iterate a cascade process on it many times keeping the property.

For $p \geq q+2$, we will denote by $\vec{\alpha}[p, q]$ the 1-deviation, where a vertex leaves a group of size $q+1$ for a group of size $p-1$ (valid as $p \geq q+2$). Let $\vec{\alpha}^1[p, q, r] = \sum_{i=0}^{r-1} \vec{\alpha}[p-i, q+i]$ denote a cascade of r such 1-deviations (we need $p-r+1 \geq q+r+2$ in order it is valid). The coordinates of $\vec{\alpha}^1[p, q, r]$ are given in the following Claim 19.

► **Claim 19.** For $p-r \geq q+r+1$, $\vec{\alpha}^1[p, q, r] = \sum_{i=0}^{r-1} \vec{\alpha}[p-i, q+i]$ has only 4 non zero coordinates namely $\alpha_p^1 = 1$, $\alpha_{p-r}^1 = -1$, $\alpha_{q+r}^1 = -1$, and $\alpha_q^1 = 1$.

► **Claim 20.** For $3 \leq t \leq \frac{p+1}{2}$, the coordinates of the sequence $\vec{\zeta}^2[p] = \vec{\delta}^2[p, t-2] + \vec{\alpha}^1[p-1, p-2t-1, t-2] + \vec{\alpha}^1[p-1, p-2t, 2] + \vec{\alpha}^1[p-t+2, p-2t+1, 1] + \vec{\alpha}^1[p-t, p-t-3, 1]$ satisfy: $\zeta_p^2 = 1$, $\zeta_{p-2}^2 = -1$, $\zeta_{p-3}^2 = -1$, $\zeta_{p-t}^2 = 1$, $\zeta_{p-t-1}^2 = 1$, $\zeta_{p-2t+2}^2 = -1$, $\zeta_{p-2t+1}^2 = -1$, $\zeta_{p-2t-1}^2 = 1$ (see Table 6). Furthermore this cascade is still 4-balanced.

Proof. By Claim 19, we have the following coordinates:

- for $\vec{\alpha}^1[p-1, p-2t-1, t-2]$, $\alpha_{p-1}^1 = 1$, $\alpha_{p-t+1}^1 = -1$, $\alpha_{p-t-3}^1 = -1$, $\alpha_{p-2t-1}^1 = 1$;
- for $\vec{\alpha}^1[p-1, p-2t, 2]$, $\alpha_{p-1}^1 = 1$, $\alpha_{p-3}^1 = -1$, $\alpha_{p-2t+2}^1 = -1$, $\alpha_{p-2t}^1 = 1$;
- for $\vec{\alpha}^1[p-t+2, p-2t+1, 1]$, $\alpha_{p-t+2}^1 = 1$, $\alpha_{p-t+1}^1 = -1$, $\alpha_{p-2t+2}^1 = -1$, $\alpha_{p-2t+1}^1 = 1$;
- for $\vec{\alpha}^1[p-t, p-t-3, 1]$, $\alpha_{p-t}^1 = 1$, $\alpha_{p-t-1}^1 = -1$, $\alpha_{p-t-2}^1 = -1$, $\alpha_{p-t-3}^1 = 1$.

XX:18 How long does it take for all users in a social network to choose their communities?

...	ζ_p^2	ζ_{p-1}^2	ζ_{p-2}^2	ζ_{p-3}^2	...	ζ_{p-t}^2	ζ_{p-t-1}^2	...	ζ_{p-2t+2}^2	ζ_{p-2t+1}^2	ζ_{p-2t}^2	ζ_{p-2t-1}^2	...
...0	1	0	-1	-1	0...0	1	1	0...0	-1	-1	0	1	0...

■ **Table 6** Difference vector $\zeta^2[p]$.

Therefore, using these values and the values of the coordinates of δ_j^2 given in claim 18, we get $\zeta_p^2 = 1$, $\zeta_{p-1}^2 = -2 + 1 + 1 = 0$, $\zeta_{p-2}^2 = -1$, $\zeta_{p-3}^2 = 0 - 1 = -1$, $\zeta_{p-t+2}^2 = -1 + 1 = 0$, $\zeta_{p-t+1}^2 = 2 - 1 - 1 = 0$, $\zeta_{p-t}^2 = 0 + 1 = 1$, $\zeta_{p-t-1}^2 = 2 - 1 = 1$, $\zeta_{p-t-2}^2 = 1 - 1 = 0$, $\zeta_{p-t-3}^2 = 0 - 1 + 1 = 0$, $\zeta_{p-2t+2}^2 = 1 - 1 - 1 = -1$, $\zeta_{p-2t+1}^2 = -2 + 1 = -1$, $\zeta_{p-2t}^2 = -1 + 1 = 0$, $\zeta_{p-2t-1}^2 = 0 + 1$.

To prove that $\vec{\zeta}^2[p]$ is 4-balanced, apply Lemma 17 with $\vec{\Phi}^1 = \vec{\delta}^2[p, t-2]$ and $\vec{\Phi}^2 = \vec{\alpha}^1[p-1, p-2t-1, t-2] + \vec{\alpha}^1[p-1, p-2t, 1] + \vec{\alpha}^1[p-t+2, p-2t+1, 1] + \vec{\alpha}^1[p-t, p-t-3, 1]$. We have that $h_1 = 4$ and furthermore all the coefficients of $\vec{\Phi}^1$ are greater than -2 and $\vec{\Phi}^2$ is 2-balanced. Hence, $\vec{\zeta}^2[p]$ is $\max(4, 2+2) = 4$ -balanced. \diamond

Table 7 shows an example with $t = 7$.

	...	p	p-1	...				p-7	p-8	...				p-15				...
$\delta^2[p, 5]$	0	1	-2	-1	0	0	-1	2	0	2	1	0	0	1	-2	-1	0	
$+\alpha[p-1,p-15,5]$	0	0	1	0	0	0	0	-1	0	0	0	-1	0	0	0	0	1	0
$+\alpha[p-1,p-14,2]$	0	0	1	0	-1	0	0	0	0	0	0	-1	0	-1	0	1	0	0
$+\alpha[p-5,p-13,1]$	0	0	0	0	0	0	1	-1	0	0	0	0	0	-1	1	0	0	0
$+\alpha[p-7,p-10,1]$	0	0	0	0	0	0	0	0	1	-1	-1	1	0	0	0	0	0	0
$= \overrightarrow{\zeta}^2[p]$	0	1	0	-1	-1	0	0	0	1	1	0	0	0	-1	-1	0	1	0

■ **Table 7** Computation of $\zeta^2[p]$ with $t = 7$.

► **Definition 21. Nice Property:** Let $k \geq 2$ be a positive integer. We will say the sequence $\vec{\zeta}^k[p]$ has the *Nice Property*, if there exist 3 integers $a(k)$, $b(k)$, and $s(k)$ satisfying $1 < a(k) < b(k) < 2a(k)$ and $b(k) < s(k) - 1 < p/2$ and such that all coordinates of $\vec{\zeta}^k$ are null except for:

- $\zeta_p^k = \zeta_{p+1-2s(k)}^k = 1$,
- $\zeta_{p-a(k)}^k = \zeta_{p-b(k)}^k = \zeta_{p+1-2s(k)+b(k)}^k = \zeta_{p+1-2s(k)+a(k)}^k = -1$, and
- $\zeta_{p+1-s(k)}^k = \zeta_{p-s(k)}^k = 1$.

We note the symmetry of the coordinates, as for any j , $\zeta_{p-j}^k = \zeta_{p+1-2s(k)+j}^k$. As an example, the sequence $\vec{\zeta}^2[p]$ satisfies the Nice Property with $a(2) = 2$, $b(2) = 3$ and $s(2) = t + 1$ and is 4-balanced. Now we will show how starting with a sequence $\vec{\zeta}^k[p]$ satisfying the Nice Property we can construct a sequence $\vec{\zeta}^{k+1}[p]$ having still the Nice Property.

► **Claim 22. Main construction:** Let $\vec{\zeta}^k[p]$ be a sequence satisfying the Nice Property with parameters $a(k), b(k), s(k)$. Then, we can construct a sequence $\vec{\zeta}^{k+1}[p]$ satisfying the following properties:

- $\vec{\zeta}^{k+1}[p]$ satisfies the Nice Property with parameters
 - $a(k+1) = b(k)$,
 - $b(k+1) = b(k) + a(k)$,

- $s(k+1) = s(k) + a(k)r(k)/2$, where $r(k)$ is the greatest even integer such that $r(k)a(k) + b(k) < s(k) - 1$;
- if $\vec{\zeta}^k[p]$ is $h(k)$ -balanced, then $\vec{\zeta}^{k+1}[p]$ is $(h(k) + 1)$ -balanced;
- $\vec{\zeta}^{k+1}[p]$ contains $r(k) + 1$ sequences $\vec{\zeta}^k[p]$.

Proof. We will first do a cascade of $\vec{\zeta}^k[p]$, but we will take values of the parameters differing by a multiple of $a(k)$ in order for some of the coordinates to cancel. More precisely let us define $\vec{\Psi}^r = \sum_{j=0}^r \vec{\zeta}^k[p - ja(k)]$. Using the values of Definition 21, we get the following values for the non zero coordinates:

- (1) $\psi_p^r = 1$; $\psi_{p-ja(k)}^r = -1 + 1 = 0$ for $0 < j \leq r$ (cancellation phenomenon); $\psi_{p-(r+1)a(k)}^r = -1$;
- (2) $\psi_{p+1-2s(k)+a(k)}^r = -1$; $\psi_{p+1-2s(k)-(j-1)a(k)}^r = 1 - 1 = 0$ for $0 < j \leq r$ (cancellation); $\psi_{p+1-2s(k)-ra(k)}^r = 1$;
- (3) for $0 \leq j \leq r$, $\psi_{p-b(k)-ja(k)}^r = -1$;
- (4) for $0 \leq j \leq r$, $\psi_{p+1-2s(k)+b(k)-ja(k)}^r = -1$;
- (5) for $0 \leq j \leq r$, $\psi_{p+1-s(k)-ja(k)}^r = \psi_{p-s(k)-ja(k)}^r = 1$.

All the indices of the coordinates are different as $a(k) < b(k) < 2a(k)$ and as soon as we choose r even and nonzero such that $p - b(k) - ra(k) > p + 1 - s(k)$ (that is equivalent to $ra(k) + b(k) < s(k) - 1$). Let us denote $a(k+1) = b(k)$, $b(k+1) = b(k) + a(k)$ and $s(k+1) = s(k) + a(k)r/2$. Then $\vec{\Psi}^r$ has already part of the Nice Property for $k+1$. Indeed we have:

- $\psi_p^r = 1$ by (1) and $\psi_{p+1-2s(k+1)}^r = 1$ by (2) with $j = r$ (as $2s(k+1) = 2s(k) + ra(k)$);
- $\psi_{p-a(k+1)}^r = \psi_{p-b(k)}^r = -1$, $\psi_{p-b(k+1)}^r = \psi_{p-b(k)-a(k)}^r = -1$ by (3) with $j = 0, 1$;
- $\psi_{p+1-2s(k+1)+b(k+1)}^r = -1$, $\psi_{p+1-2s(k+1)+a(k+1)}^r = -1$ by (4) with $j = r-1, r$;
- $\psi_{p+1-s(k+1)}^r = \psi_{p-s(k+1)}^r = 1$ by (5) with $j = r/2$.

The remaining non zero coordinates are in number $4r$: firstly there are r values -1 , namely $\psi_{p-b(k)-ja(k)}^r = -1$, for $2 \leq j \leq r$, and $\psi_{p-(r+1)a(k)}^r = -1$; then there are $2r$ values 1 , namely $\psi_{p+1-s(k+1)}^r = \psi_{p-s(k+1)}^r = 1$, for $j \neq r/2$ and finally r values -1 , namely $\psi_{p+1-2s(k)+a(k)}^r = -1$ and $\psi_{p+1-2s(k)+b(k)-ja(k)}^r = -1$, for $0 \leq j \leq r-2$. These values are disposed in a very symmetric way and can be written: for the values -1 , in the form $\psi_{p-x_m}^r$ and $\psi_{p+1-2s(k+1)+x_m}^r$; and for the values 1 , in the form $\psi_{p-y_m}^r$ and $\psi_{p+1-2s(k+1)+y_m}^r$ with $x_m < y_m$ ($0 \leq m \leq r-1$). Furthermore, these r quadruples of values can be canceled by adding to $\vec{\Psi}^r$ the r sequences $\vec{\alpha}^1[p - x_m, p + 1 - 2s(k+1) + x_m, y_m - x_m]$.

We claim that the sequence so obtained, with partition vector $\vec{\Psi}^r + \sum_{m=0}^{r-1} \vec{\alpha}^1[p - x_m, p + 1 - 2s(k+1) + x_m, y_m - x_m]$, satisfies the Nice Property with parameters $a(k+1), b(k+1)$ and $s(k+1)$. Indeed $a(k+1) = b(k) < b(k) + a(k) = b(k+1)$, $b(k+1) = b(k) + a(k) < b(k) + b(k) = 2a(k+1)$ and $b(k+1) = b(k) + a(k) < s(k) - 1 + a(k) \leq s(k+1) - 1$ as $r \geq 2$. We also have to ensure in the computations that p is chosen so that $p \geq 2s(k) - 1$. In order to get the maximum number of deviations we will consider this sequence for the largest possible even integer r satisfying $ra(k) + b(k) < s(k) - 1$, denoted $r(k)$ and we will denote the sequence for this $r(k)$ by $\vec{\zeta}^{k+1}[p]$.

We now prove that $\vec{\zeta}^{k+1}[p]$ is $h(k) + 1$ -balanced. We first prove by induction that $\vec{\Psi}^r$ is $(h(k) + 1)$ -balanced. That is true for $r = 0$ as $\vec{\zeta}^k[p]$ is $h(k)$ -balanced. Then suppose it is true for some r ; we apply Lemma 17 with $\vec{\Phi}^1 = \vec{\Psi}^r$ and $\vec{\Phi}^2 = \vec{\zeta}^k[p - (r+1)a(k)]$. We have that $h_1 = h(k) + 1$ by induction hypothesis and furthermore all the coefficients of $\vec{\Phi}^1$ are greater than -1 ; furthermore $\vec{\Phi}^2$ is $h(k)$ -balanced and so $\vec{\Psi}^{r+1}$ is $(\max(h(k) + 1, h(k) + 1)) =$

706 $h(k) + 1$ -balanced. Then when we add an $\vec{\alpha}^1[p - x_m, p + 1 - 2s(k + 1) + x_m, y_m - x_m]$
 707 which is 1-balanced we still get an $(\max(h(k) + 1, 1 + 1) = h(k) + 1)$ -balanced sequence.
 708 Finally, by construction, we get that $\vec{\zeta}^{k+1}[p]$ contains $r(k) + 1$ sequences $\vec{\zeta}^k[p]$. \diamond

709 **End of the proof of Theorem 14.**

710 At this stage we have built a sequence $\vec{\zeta}^2[p]$ which satisfies the Nice Property with
 711 $a(2) = 2$, $b(2) = 3$ and $s(2) = t + 1$ and is $h(2) = 4$ -balanced. Furthermore, it contains $t(t - 2)$
 712 4-deviations. See Claim 20. Then, for some well-chosen K (to be defined later) we can apply
 713 $K - 2$ times the main construction (Claim 22) to construct a sequence $\vec{\zeta}^K[p]$ which satisfies
 714 the Nice Property with parameters $a(K)$, $b(K)$ and $s(K)$ and is $h(K)$ -balanced.

We have $a(k) = b(k - 1)$, $b(k) = b(k - 1) + a(k - 1) = b(k - 1) + b(k - 2)$ and we recognize the Fibonacci recurrence relation. The k^{th} Fibonacci number $F(k)$ is denoted as follows:

$$F(k) = \frac{1}{\sqrt{5}} \left(\left(\frac{1 + \sqrt{5}}{2} \right)^k - \left(\frac{1 - \sqrt{5}}{2} \right)^k \right).$$

715 Then, as $a(2) = 2 = F(3)$ and $b(2) = 3 = F(4)$, we get $a(K) = F(K + 1)$ and $b(K) =$
 716 $F(K + 2)$. In fact in what follows we will use only that $a(K) \leq 2^{K-1}$ and $b(K) \leq 2^K$.
 717 We have $s(k + 1) = s(k) + a(k)r(k)/2$; but $a(k)r(k) < s(k - 1) - b(k) < s(k - 1)$ and so
 718 $s(k + 1) < (3/2) \times s(k)$ and $s(K) < s(2)(3/2)^{K-2} = (t + 1)(3/2)^{K-2}$.

719 Recall that we should have $p \geq 2s(K) - 1$ so we choose $p = 2s(K)$. Furthermore by
 720 induction we have that $h(K) = K + 2$. So we need to start with a partition containing at
 721 least $K + 2$ groups of each size i , $1 \leq i \leq p$. It is easy to obtain such a starting partition
 722 from the initial partition — which consists of n groups of size 1 — by doing a sequence of
 723 1-deviation of size $(K - 2)p(p + 1)/2$; indeed we can create a group of any size i with $(i - 1)$
 724 1-deviations. Therefore, we will take $n = (K - 2)p(p + 1)/2 \leq (K - 2)s(K)(2s(K) + 1)$.
 725 Using the inequality $s(K) < (t + 1)(3/2)^{K-2}$ we get that

$$726 \quad n = \mathcal{O}(t^2 K (3/2)^{2K}). \quad (1)$$

727 On the other hand we have to lower bound the number of deviations. By construction
 728 $\vec{\zeta}^{k+1}[p]$ contains $r(k) + 1$ sequences $\vec{\zeta}^k[p]$ and so, contains $t(t - 2) \prod_{k=2}^{K-1} (r(k) + 1)$ 4-
 729 deviations, as $\vec{\zeta}^2[p]$ contains $t(t - 2)$ 4-deviations. Recall that $r(k)$ is the greatest even
 730 integer r such that $ra(k) + b(k) < s(k) - 1$ and so $r(k) \geq \lfloor \frac{s(k)-1-b(k)}{a(k)} \rfloor - 1$. Using the
 731 fact that $b(k) + 1 \leq 2a(k)$ and $s(k) > s(2) - 1 = t$, and $a(k) \leq a(K) < 2^{K-1}$ we get
 732 $r(k) \geq \frac{t}{2^{K-1}} - 3$. Then $\prod_{k=2}^{K-1} (r(k) + 1) \geq (\frac{t}{2^{K-1}} - 2)^{K-2}$ and the number D of deviations
 733 satisfies:

$$734 \quad D = \Omega(t^2 (\frac{t}{2^{K-1}} - 2)^{K-2}). \quad (2)$$

735 We have now to choose K as a function of t . In order for the number of deviations as given
 736 by Equation 2 to increase we need that 2^{K-1} is small compared to t that is $K \ll \log_2(t)$.
 737 However in view of Equation 1 we want to choose K the largest possible. Therefore, a good
 738 choice is $K = 1/2(\log_2(t))$.

739 In that case, we get by Equation 1 that $n = \mathcal{O}(t^2 \log_2(t) (3/2)^{\log_2(t)})$, or equivalently
 740 $\log_2(n) = \mathcal{O}(2 \log_2(t) + \log_2(\log_2(t) + \log_2(t)(\log_2(3) - \log_2(2))))$. Using $\log_2(3) - \log_2(2) >$
 741 0.585 and the fact that for t large enough $\log_2(\log_2(t)) < 0.014 \log_2(t)$ we get $\log_2(n) =$
 742 $\mathcal{O}(2.6 \log_2(t))$, that is $n = \mathcal{O}(t^{2.6})$. On the other hand we get by Equation 2: $D =$
 743 $\Omega((t^{1/2})^{1/2 \log_2(t)}) = \Omega(t^{1/4 \log_2(t)})$ and so $D = \Omega(n^{c \log_2(n)})$ with $c = \frac{1}{4 \times (2.6)^2} \simeq 1/27$, thereby
 744 proving Theorem 14.

References

- 1 C. Ballester. NP-completeness in hedonic games. *Games and Economic Behavior*, 49(1):1–30, 2004.
- 2 J. A. Bondy and U. S. R. Murty. *Graph theory*. Grad. Texts in Math., 2008.
- 3 T. Brylawski. The lattice of integer partitions. *Discrete Mathematics*, 6(3):201 – 219, 1973. doi:10.1016/0012-365X(73)90094-0.
- 4 I. Chatzigiannakis, C. Koninis, P. N. Panagopoulou, and P. G. Spirakis. Distributed game-theoretic vertex coloring. In *OPODIS’10*, pages 103–118, 2010.
- 5 J. Chen, R. Niedermeier, and P. Skowron. Stable marriage with multi-modal preferences. In *EC*, 2018. to appear.
- 6 G. Ducoffe. *Propriétés métriques des grands graphes*. PhD thesis, Université Côte d’Azur, December 2016.
- 7 B. Escoffier, L. Gourvès, and J. Monnot. Strategic coloring of a graph. *Internet Mathematics*, 8(4):424–455, 2012.
- 8 M. Flammini, G. Monaco, and Q. Zhang. Strategyproof mechanisms for additively separable hedonic games and fractional hedonic games. In *WAOA*, pages 301–316, 2017.
- 9 C. Greene and D. J. Kleitman. Longest chains in the lattice of integer partitions ordered by majorization. *European Journal of Combinatorics*, 7(1):1–10, jan 1986. URL: <http://dl.acm.org/citation.cfm?id=11002.11003>.
- 10 J. Hajduková. Coalition formation games: A survey. *International Game Theory Review*, 8(04):613–641, 2006.
- 11 G. H. Hardy and E. M. Wright. *An introduction to the theory of numbers*. Oxford University Press, 1979.
- 12 M. Hoefer and W. Jiamjitrak. On proportional allocation in hedonic games. In *SAGT*, pages 307–319. Springer, 2017.
- 13 D. S. Johnson, C. H. Papadimitriou, and M. Yannakakis. How easy is local search? *Journal of computer and system sciences*, 37(1):79–100, 1988.
- 14 J. Kleinberg and K. Ligett. Information-sharing in social networks. *Games and Economic Behavior*, 82:702–716, 2013.
- 15 M. Mnich and I. Schlotter. Stable marriage with covering constraints—a complete computational trichotomy. In *SAGT*, pages 320–332. Springer, 2017.
- 16 K. Ohta, N. Barrot, A. Ismaili, Y. Sakurai, and M. Yokoo. Core stability in hedonic games among friends and enemies: impact of neutrals. In *IJCAI*, 2017.
- 17 I. Olkin and A. W. Marshall. *Inequalities: theory of majorization and its applications*, volume 143. Academic press, 2016.
- 18 P. N. Panagopoulou and P. G. Spirakis. A game theoretic approach for efficient graph coloring. In *ISAAC’08*, pages 183–195, 2008.

5.2 Appendix: Proof for the case $k = 3$ (Theorem 13)

The proof uses the same idea that for $k = 4$ but is simpler as we can only do a limited iteration of cascades, which use a lot of 3-deviations $\gamma[p]$ (the only 3-deviation which does not increase the global utility).

Definition of $\gamma[p]$: Consider a partition P containing at least three groups of size $p - 1$ and a group of size $p - 3$. In the 3-deviation $\gamma[p]$ one vertex in each of the 3 groups of size $p - 1$ moves to the group of size $p - 3$ to form a new group of size p . The example given at the beginning of section 5 corresponds to the case $p = 4$. So, after this 3-deviation we get a new partition P' with one more group of size p , 3 less groups of size $p - 1$, 3 more groups of size $p - 2$ and one less group of size $p - 3$. This is expressed by the coordinates of the associated difference vector (where we omit the bracket $[p]$ for ease of reading). Note that such a deviation is valid only if there are 3 groups of size $p - 1$ and one group of size $p - 3$.

Difference vector $\vec{\gamma}[p]$: The difference vector $\vec{\gamma}[p]$ has the following coordinates: $\gamma_p = 1, \gamma_{p-1} = -3, \gamma_{p-2} = 3, \gamma_{p-3} = -1$ and $\gamma_j = 0$ for all other values of j . See Table 8.

...	γ_p	γ_{p-1}	γ_{p-2}	γ_{p-3}	...
...0	1	-3	3	1	0...

■ **Table 8** Difference vector of $\gamma[p]$.

The cascade $\vec{\gamma}^1[p, t]$: Like for $k = 4$, we do a cascade consisting of t 3-deviations $\gamma[p], \gamma[p - 1], \gamma[p - 2], \dots, \gamma[p - t + 1]$. We will denote this cascade by its difference vector $\vec{\gamma}^1[p, t] = \sum_{i=0}^{t-1} \vec{\gamma}[p - i]$. We will determine the coordinates of $\vec{\gamma}^1[p, t]$ in Claim 23, but first let us see how it works on a basic example.

p	13	12	11	10	9	8	7	6	5	4	3	2	1
$\vec{\gamma}[13]$	1	-3	3	-1									
$\vec{\gamma}[12]$		1	-3	3	-1								
$\vec{\gamma}[11]$			1	-3	3	-1							
$\vec{\gamma}[10]$				1	-3	3	-1						
$\vec{\gamma}[9]$					1	-3	3	-1					
$\vec{\gamma}[8]$						1	-3	3	-1				
$\gamma^1[13, 6]$	1	-2	1	0	0	0	-1	2	-1	0	0	0	0

■ **Table 9** Computation of $\gamma^1[13, 6]$.

Basic example: We consider the partition P^0 consisting of $g \geq 3$ groups of each size i for $1 \leq i \leq 12$. The partition vector associated $\vec{\lambda}(P^0)$ satisfies $\lambda_i^0 = g$ for $1 \leq i \leq 12$ and $\lambda_j^0 = 0$ for $j \geq 13$. In the example of Table 9, we choose $t = 6$ and so we do successively $\vec{\gamma}[13 - i]$. The non zero coordinates of $\vec{\gamma}[13 - i]$ are indicated in the corresponding line. The j th coordinate of $\vec{\gamma}^1[13, r]$ is obtained by summing the numbers of the first r lines and column j . The coordinate indicates the number of groups of size j created (if positive) or deleted (if negative). If we look at the cascade obtained after 4 3-deviations we note a remarkable phenomenon of cancellation as the coordinate γ_{10}^1 of this cascade equals 0. That comes from the fact that we have successively deleted a group of size 10 with $\vec{\gamma}[13]$, then created 3 new groups with $\vec{\gamma}[12]$, then deleted 3 groups with $\vec{\gamma}[11]$ and finally created a new group with $\vec{\gamma}[10]$. This cancellation stays for all the other deviations. Similarly, the coordinate γ_9^1 of the cascade $\vec{\gamma}^1[13, 5]$ equals 0 and so on. In fact as indicated in the

812 Claim 23 the vector $\vec{\gamma}^1[p, t]$ has only 6 non zero coordinates.

813 ► **Claim 23.** For $3 \leq t \leq p - 3$, the coordinates of the cascade $\vec{\gamma}^1[p, t] = \sum_{i=0}^{t-1} \vec{\gamma}[p - i]$
 814 satisfy: $\gamma_p^1 = 1$, $\gamma_{p-1}^1 = -2$, $\gamma_{p-2}^1 = 1$, $\gamma_{p-t}^1 = -1$, $\gamma_{p-t-1}^1 = 2$, $\gamma_{p-t-2}^1 = -1$ and $\gamma_j^1 = 0$ for
 815 all the others j . Furthermore, this cascade is 3-balanced.

...	γ_p^1	γ_{p-1}^1	γ_{p-2}^1		...	γ_{p-t}^1	γ_{p-t-1}^1	γ_{p-t-2}^1	...
...0	1	-2	1	0...0	-1	2	-1	0...	

■ **Table 10** Difference vector $\gamma^1[p, t]$.

816 *Proof.* We have $\gamma_j^1 = \sum_{i=0}^{t-1} \gamma_j[p - i]$. For a given j , $\gamma_j[p - i] = 0$ except for the following
 817 values of i such that $0 \leq i \leq t - 1$: $i = p - j$ where $\gamma_j[j] = 1$; $i = p - j - 1$ where
 818 $\gamma_j[j + 1] = -3$; $i = p - j - 2$ where $\gamma_j[j + 2] = 3$; $i = p - j - 3$ where $\gamma_j[j + 3] = -1$ (in
 819 the table it corresponds to the consecutive non zero values in a column which are at most
 820 4). Therefore, for $j > p$: $\gamma_j^1 = 0$; $\gamma_p^1 = 1$; $\gamma_{p-1}^1 = -3 + 1 = -2$; $\gamma_{p-2}^1 = 3 - 3 + 1 = 1$; for
 821 $p - 3 \geq j \geq p - t + 1$, $\gamma_{p-j}^1 = -1 + 3 - 3 + 1 = 0$; $\gamma_{p-t}^1 = -1 + 3 - 3 = -1$; $\gamma_{p-t-1}^1 = -1 + 3 = 2$;
 822 $\gamma_{p-t-2}^1 = -1$ and, for $j < p - t - 2$, $\gamma_j^1 = 0$. Finally we note that the coordinates of any
 823 subsequence of the cascade, *i.e.* the coordinates of $\vec{\gamma}^1[p, r]$, are all at least -1 except γ_{p-1}^1 :
 824 which is -3 when $r = 1$ and then -2 , thereby proving that the cascade is 3-balanced. \diamond

825 In our example we can do the cascade till $t = 10$ and in general we can do it till
 826 $t = p - 3$. However, it is better to choose a t smaller (we will see after that a good value
 827 is such that $p - 3t > 0$) and repeat the previous cascade but shifted, and so, to do a
 828 cascade of cascades. More precisely we do now the following sequence of $t - 1$ cascades
 829 $\vec{\gamma}^2[p, t - 1] = \sum_{i=0}^{t-2} \vec{\gamma}^1[p - i, t]$. Altogether we have a sequence of $t(t - 1)$ 3-deviations.

p	13	12	11	10	9	8	7	6	5	4	3	2	1
$\gamma^1[13, 4]$	1	-2	1	0	-1	2	-1						
$\gamma^1[12, 4]$		1	-2	1	0	-1	2	-1					
$\gamma^1[11, 4]$			1	-2	1	0	-1	2	-1				
$\gamma^2[13, 3]$	1	-1	0	-1	0	1	0	1	-1				
$\gamma^2[12, 3]$		1	-1	0	-1	0	1	0	1	-1			
$\gamma^3[13, 2]$	1	0	-1	-1	-1	1	1	1	0	-1			
$\alpha^1[11, 4, 3]$	0	0	1	0	0	-1	-1	0	0	1			
$\zeta^3[13, 2]$	1	0	0	-1	-1	0	0	1	0	0			
$\zeta^3[12, 2]$		1	0	0	-1	-1	0	0	1	0	0		
$\zeta^3[11, 2]$			1	0	0	-1	-1	0	0	1	0	0	
$\zeta^3[10, 2]$				1	0	0	-1	-1	0	0	1	0	0
$\zeta^4[13, 4]$	1	1	1	0	-2	-2	-2	0	1	1	1	0	0

■ **Table 11** Example of cascade of cascades.

830 In the example (see Table 11), we choose $t = 4$ and so after having done $\vec{\gamma}^1[13, 4]$ we do
 831 $\vec{\gamma}^1[12, 4]$ and $\vec{\gamma}^1[11, 4]$. We can see again a phenomenon of cancellation as the coordinate
 832 γ_{11}^2 of this cascade of cascade equals 0. That comes from the fact that we have successively
 833 created a group of size 11 with $\vec{\gamma}^1[13, 4]$, then deleted 2 groups with $\vec{\gamma}^1[12, 4]$ and finally
 834 created a new group with $\vec{\gamma}^1[11, 4]$. Similarly the coordinate γ_7^2 of this cascade of cascade
 835 equals 0. This cancellation stays for all the other deviations. In the general case there are a
 836 lot of cancellations and in fact, as shown in the next Claim 24, $\vec{\gamma}^2[p, t - 1]$ has only 6 non
 837 zero coordinates.

► **Claim 24.** For $3 \leq t \leq \frac{p-1}{2}$, the coordinates of the cascade $\vec{\gamma}^2[p, t-1] = \sum_{i=0}^{t-2} \vec{\gamma}^1[p-i, t]$ satisfy: $\gamma_p^2 = 1$, $\gamma_{p-1}^2 = -1$, $\gamma_{p-t+1}^2 = -1$, $\gamma_{p-t-1}^2 = 1$, $\gamma_{p-2t+1}^2 = 1$, $\gamma_{p-2t}^2 = -1$ and $\gamma_j^2 = 0$ for all the others j . Furthermore this cascade is 3-balanced.

...	γ_p^2	γ_{p-1}^2	...	γ_{p-t+1}^2		γ_{p-t-1}^2	...	γ_{p-2t+1}^2	γ_{p-2t}^2	...
0	1	-1	0	-1	0	1	0	1	-1	0

■ **Table 12** Difference vector $\delta^2[p, t-2]$.

Proof. We have $\gamma_j^2 = \sum_{i=0}^{t-2} \gamma_j^1[p-i, t]$. Using the values of $\gamma_j^1[p-i, t]$ given in Claim 23, we get that: for $j > p$, $\gamma_j^2 = 0$; $\gamma_p^2 = 1$; $\gamma_{p-1}^2 = 1 - 2 = -1$; for $p-2 \geq j \geq p-t+2$, $\gamma_j^2 = 1 - 2 + 1 = 0$; $\gamma_{p-t+1}^2 = 1 - 2 = -1$; $\gamma_{p-t}^2 = -1 + 0 + 1 = 0$; $\gamma_{p-t-1}^2 = 2 - 1 = 1$; for $p-t-2 \geq j \geq p-2t+2$, $\gamma_j^2 = -1 + 2 - 1 = 0$; $\gamma_{p-2t+1}^2 = -1 + 2 = 1$; $\gamma_{p-2t}^2 = -1$ and for $j < p-2t$, $\gamma_j^2 = 0$.

Here again we can see that after any number r of 3-deviations the coordinates of the sequence are always at least greater than or equal to -3 . Indeed the -3 appears only after the first 3-deviation for groups of size $p-1$; otherwise, when a -3 appears it is after a $+3$ and so all the other coordinates are in fact at least -2 . Thus the cascade is 3-balanced. \diamond

We can now repeat the cascade of cascades $\vec{\gamma}^2[p, t-1]$ but shifted. More precisely we do now the following sequence of $t-2$ cascades $\vec{\gamma}^3[p, t-2] = \sum_{i=0}^{t-3} \vec{\gamma}^2[p-i, t-1]$. Altogether we have a sequence of $t(t-1)(t-2)$ 3-deviations. In the example (see Table 11), after having done $\vec{\gamma}^2[13, 3]$ we do $\vec{\gamma}^2[12, 3]$. We can see again a phenomenon of cancellation as the coordinate γ_{12}^3 of this cascade of cascades equals 0. That comes from the fact that we have successively deleted a group of size 12 after $\vec{\gamma}^2[13, 3]$, then created a new group after $\vec{\gamma}^2[12, 3]$. Similarly the coordinate γ_5^3 of this cascade of cascade equals 0. Here we have also deleted one group of size 5, then created one such group. This cancellation stays for all the other deviations. In the general case there are a lot of cancellations and in fact as shown in the next Claim 25, $\vec{\gamma}^2[p, t-1]$ has only 8 non zero coordinates.

► **Claim 25.** For $3 \leq t \leq \frac{p+2}{3}$, the coordinates of the cascade $\vec{\gamma}^3[p, t-2] = \sum_{i=0}^{t-3} \vec{\gamma}^2[p-i, t-1]$ satisfy: $\gamma_p^3 = 1$, $\gamma_{p-t+2}^3 = \gamma_{p-t+1}^3 = \gamma_{p-t}^3 = -1$, $\gamma_{p-2t+3}^3 = \gamma_{p-2t+2}^3 = \gamma_{p-2t+1}^3 = 1$, $\gamma_{p-3t+3}^3 = -1$, and $\gamma_j^3 = 0$ for all the others j . Furthermore this cascade is 4-balanced

...	γ_p^3	...	γ_{p-t+2}^3	γ_{p-t+1}^3	γ_{p-t}^3	...	γ_{p-2t+3}^3	γ_{p-2t+2}^3	γ_{p-2t+1}^3	...	γ_{p-3t+3}^3	...
0	1	0	-1	-1	-1	0	1	1	1	0	-1	0

■ **Table 13** Difference vector $\gamma^3[p, t-2]$.

Proof. We have $\gamma_j^3 = \sum_{i=0}^{t-3} \gamma_j^2[p-i, t-1]$. Using the values of $\gamma_j^2[p-i, t-1]$ given in Claim 24, we get that: for $j > p$, $\gamma_j^3 = 0$; $\gamma_p^3 = 1$; for $p-1 \geq j \geq p-t+3$, $\gamma_j^3 = -1 + 1 = 0$; $\gamma_{p-t+2}^3 = 0 - 1 = -1$; $\gamma_{p-t+1}^3 = -1 + 0 = -1$; $\gamma_{p-t}^3 = 0 - 1 = -1$; for $p-t-1 \geq j \geq p-2t+4$, $\gamma_j^3 = 1 + 0 - 1 = 0$; $\gamma_{p-2t+3}^3 = 1 + 0 = 1$; $\gamma_{p-2t+2}^3 = 0 + 1 = 1$; $\gamma_{p-2t+1}^3 = 1 + 0 = 1$; for $p-2t \geq j \geq p-3t+4$, $\gamma_j^3 = -1 + 1 = 0$; $\gamma_{p-3t+3}^3 = 0 - 1 = -1$; and for $j < p-3t+3$, $\gamma_j^3 = 0$.

Here again a careful but tedious analysis of all the subsequence of deviations indicate that all their coordinates are at least -3 . However using Lemma 17 we can easily prove that it is 4-balanced. In fact, we will prove by induction that $\sum_{i=0}^r \vec{\gamma}^2[p-i, t-1]$ is 4-balanced for any $r \leq t-3$. That is true for $r = 0$, as $\vec{\gamma}^2[p, t-1]$ is 3-balanced. Suppose

that $\sum_{i=0}^r \vec{\gamma}^2[p-i, t-1]$ is 4-balanced for some $r \leq t-3$. We apply Lemma 17 with $\vec{\Phi}^1 = \sum_{i=0}^r \vec{\gamma}^2[p-i, t-1]$ and $\vec{\Phi}^2 = \vec{\gamma}^2[p-r-1, t-1]$. We have by induction hypothesis that $h_1 = 4$ and furthermore all the coefficients of $\vec{\Phi}^1$ are greater than -1 by Claim 24 when $r = 0$ or Claim 25 when $r > 0$ and so $\min_i \Phi_i^1 = -1$. Finally $\vec{\Phi}^2$ is 3-balanced and so $\sum_{i=0}^{r+1} \vec{\gamma}^2[p-i, t-1]$ is also $\max(4, 3+1) = 4$ -balanced and by induction $\vec{\gamma}^3[p, t-2]$ is 4-balanced. \diamond

We can now again repeat t times the cascades $\vec{\gamma}^3[p, t-2]$ creating a sequence of cascades that we call $\vec{\gamma}^4[p, t] = \sum_{i=0}^{t-1} \vec{\gamma}^3[p-i, t-2]$. This is enough to prove Theorem 13. But we will use the trick used for $k = 4$ which consists in doing after $\vec{\gamma}^3[p, t-2]$ a sequence of 1-deviations which will reduce to 4 the number of non zero coordinates. This slightly improves the ratio between the number of 3-deviations and the number of vertices. So, after having done $\vec{\gamma}^3[p, t-2]$, we apply the cascade $\vec{\alpha}^1[p-t+2, p-3t+3, t-1]$ (see Claim 19). Let $\vec{\zeta}^3[p, t-2]$ be the sequence we obtain. As shown in Claim 26, it has only 4 non zero coordinates.

► **Claim 26.** For $3 \leq t \leq p/2$ the coordinates of $\vec{\zeta}^3[p, t-2] = \vec{\gamma}^3[p, t-2] + \vec{\alpha}^1[p-t+2, p-3t+3, t-1]$ satisfy: $\zeta_p^3 = 1, \zeta_{p-t+1}^3 = \zeta_{p-t}^3 = -1, \zeta_{p-2t+1}^3 = 1$ and $\zeta_j^3 = 0$ for all the others j . Furthermore it is 4-balanced.

Proof. Compared to $\vec{\gamma}^3[p, t-2]$ only 4 coordinates have been changed and we get $\zeta_{p-t+2}^3 = \gamma_{p-t+2}^3 + \alpha^1(p-t+2) = -1+1 = 0$, $\zeta_{p-2t+3}^3 = \gamma_{p-2t+3}^3 = \zeta_{p-2t+2}^3 = 1-1 = 0$ and $\zeta_{p-3t+3}^3 = -1+1 = 0$. All the other coordinates remain the same. To prove that $\vec{\zeta}^3[p, t-2]$ is 4-balanced, we apply Lemma 17 with $\vec{\Phi}^1 = \vec{\gamma}^3[p, t-2]$ and $\vec{\Phi}^2 = \vec{\alpha}^1[p-t+2, p-3t+3, t-1]$. We have proved that $\vec{\Phi}^1$ is 4-balanced and by Claim 25 $\min_i \Phi_i^1 = -1$. Furthermore, $\vec{\Phi}^2$ is 1-balanced and so $\vec{\zeta}^3[p, t-2]$ is $\max(4, 1+1) = 4$ -balanced. \diamond

Now we can do t cascades of $\vec{\zeta}^3$ to obtain the cascade $\vec{\zeta}^4[p, t] = \sum_{i=0}^{t-1} \vec{\zeta}^3[p-i, t-2]$.

► **Claim 27.** For $3 \leq t \leq p+2/3$ the coordinates of $\vec{\zeta}^4[p, t] = \sum_{i=0}^{t-1} \vec{\zeta}^3[p-i, t-2]$ satisfy: For $p \geq j \geq p-t+2$, $\zeta_j^4 = 1$; for $p-t \geq j \geq p-2t+2$, $\zeta_j^4 = -2$; for $p-2t \geq j \geq p-3t+2$, $\zeta_j^4 = 1$; and $\zeta_j^4 = 0$ for all the others j . Furthermore it is 6-balanced.

Proof. This follows easily from the values of the coordinates of $\vec{\zeta}^3[p-i, t-2]$. Note that $\zeta_{p-t+1}^4 = -1+1 = 0$ and $\zeta_{p-2t+1}^4 = 1-1 = 0$.

We will prove by induction that $\sum_{i=0}^r \vec{\zeta}^3[p-i, t-2]$ is 6-balanced for any $r \leq t-1$. That is true for $r = 0$ as $\vec{\zeta}^3[p, t-2]$ is 4-balanced. Suppose that $\sum_{i=0}^r \vec{\zeta}^3[p-i, t-2]$ is 6-balanced for some $r \leq t-2$. We apply Lemma 17 with $\vec{\Phi}^1 = \sum_{i=0}^r \vec{\zeta}^3[p-i, t-2]$ and $\vec{\Phi}^2 = \vec{\zeta}^3[p-r-1, t-2]$. We have $h_1 = 6$ by induction hypothesis and furthermore all the coefficients of $\vec{\Phi}^1$ are greater than -2 by Claim 26 when $r = 0$ or Claim 27 when $r > 0$ and so $\min_i \Phi_i^1 = -2$. Finally $\vec{\Phi}^2$ is 4-balanced and so $\sum_{i=0}^{r+1} \vec{\zeta}^3[p-i, t-2]$ is also $\max(6, 4+2) = 6$ -balanced and by induction $\vec{\zeta}^4[p, t]$ is 6-balanced. \diamond

End of the proof of Theorem 13: The cascade $\vec{\zeta}^4[p, t]$ consists of t cascades $\vec{\zeta}^3[p-i, t-2]$ each of them consisting of a cascade $\vec{\gamma}^3[p-i, t-2]$ and $t-1$ 1-deviations. $\vec{\gamma}^3[p-i, t-2]$ itself consists of $t-2$ cascades $\vec{\gamma}^2[p-i, t-1]$ each of them consisting of $t-1$ cascades $\vec{\gamma}^1[p-i, t-1]$ each of them consisting of t 3-deviations $\vec{\gamma}[p-i]$. So the cascade $\vec{\zeta}^4[p, t]$ contains $t^2(t-1)(t-2) = \theta(t^4)$ 3-deviations (plus $t(t-1)$ 1-deviations which is negligible).

By Claim 27, it is 6-balanced and so if we choose as starting partition one with 6 groups of each size $i, 1 \leq i \leq p-1$ the cascade is valid. It is easy to obtain such a starting partition from the initial partition which consists of n groups of size 1. Indeed we can create a group

XX:26 How long does it take for all users in a social network to choose their communities?

918 of any size i ; for that we choose a specific group of size 1 and successively move with $(i - 1)$
919 1-deviations one element of $i - 1$ other groups of size 1 to form a group of size i . We do it
920 6 times for each size $i, 1 \leq i \leq p - 1$. Of course that is possible only if $n \geq 6p(p - 1)/2$.

921 Finally, note that in order that the coordinates of all cascades have a meaning we must
922 choose p and t such that $p - 3t \geq 0$. Let us choose $p = 3t$; then the number of vertices is
923 $n = 6 \sum_{i=0}^{p-1} i = 6p(p - 1)/2 = 9t(3t - 1) = \theta(t^2)$.

924 In summary we have built a cascade which contains $\theta(t^4) = \theta(n^2)$ 3-deviations and so
925 $L(3, n) = \Omega(n^2)$.