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

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

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

Jean-Claude Bermond [a]*, Augustin Chaintreau [b], Guillaume Ducoffe [c]†, Dorian Mazauric [d]‡
(a) Université Côte d'Azur, CNRS, Inria, I3S, France
(b) Columbia University in the City of New York
(c) National Institute for Research and Development in Informatics, and Research Institute and Faculty of Mathematics and Computer Science of the University of Bucharest, Bucureşti, România
(d) 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 [EGM12, KL13]. In this paper we disprove this. Specifically, we prove that for any $k \geq 4$, the maximum time of convergence is in $\Omega(|V|^{\Theta(\log |V|)})$.

Keywords: communities, social networks, integer partitions, coloring games, graphs, algorithms.

## 1 Introduction

Community formation is a fundamental problem in social network analysis. It has already been modeled in several ways, each trying to capture key aspects of the problem. The model studied in this paper has been proposed in [KL13] in order to reflect the impact of information sharing on the community formation process. Although it is a simplified model, we show that its understanding requires us to solve combinatorial problems that are surprisingly intricate. More precisely, we consider the following dynamics of formation of groups (communities) in social networks. Each group represents a set of users sharing about some information topic. We assume for simplicity that each user shares about a given topic in only one group. Therefore the groups will partition the set of users. We follow the approach of [KL13]. An important feature is the emphasis on incompatibility between some pairs of users that we will call enemies. Two enemies do not want to share information. Hence, they will necessarily belong to different groups. In the general model one considers different degrees of friendship or incompatibilities. Here we will restrict to the case where two users are

*This work has been supported by ANR program Investments for the Future under reference ANR-11-LABX-0031-01.

†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, and by a grant of Romanian Ministry of Research and Innovation CCCDI-UEFISCDI. project no. 17PCCDI/2018.

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

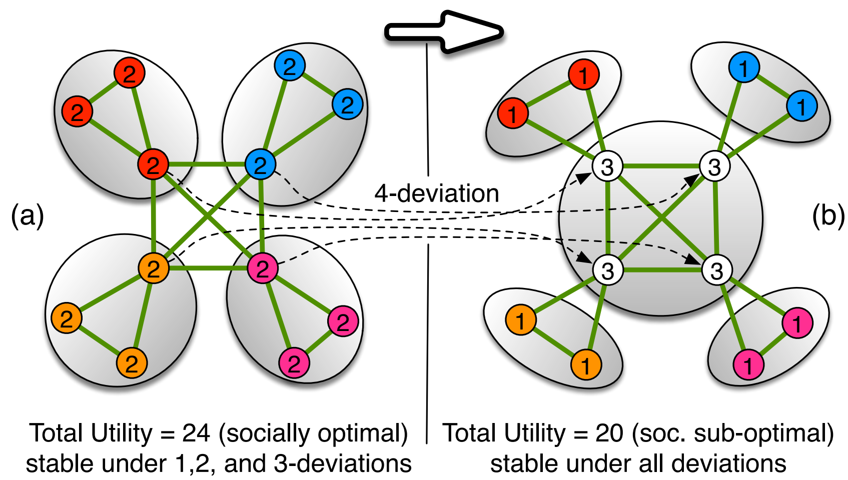


Figure 1: A friendship graph with 12 vertices (users). **(a)** 3-stable partition that is not 4-stable but it is optimal in terms of total utility. **(b)** $k$-stable partition for any $k \geq 1$ that is not optimal in terms of total utility.

either friends or enemies – as noted in [KL13], even a little beyond this case, the problem quickly becomes intractable. As an example, if we add a neutral (indifference) relation, there are instances for which there is no stability.

The social network is often modeled by the friendship graph $G^+$ where the vertices are the users and an edge represents a friendship relation. We will use this graph to present the first notions and examples. However, for the rest of the article and the proofs we will use the complementary graph, that we call the conflict graph and denote by $G^-$; here the vertices represent users and the edges represent the incompatibility relation. We assign each user a *utility* which is the number of friends in the group to which she belongs. Equivalently, the utility is the size of the group minus one, as in a group there is no pair of enemies; in [KL13] this is modeled by putting the utility as $-\infty$ when there is an enemy in the group.

In the example of Figure 1, the graph depicted is the friendship graph: the edges represent the friendship relation, and if there is no edge then, it corresponds to a pair of enemies. Figure 1(a) depicts a partition of 12 users that is composed of 4 non-empty groups each of size 3. The integers on the vertices represent the utilities of the users which are all equal to 2. This implies that the total utility is 24. Figure 1(b) depicts another partition consisting of 5 groups with one group of size 4 (where users have utility 3) and 4 groups of size 2 (where users have utility 1). Hence, the total utility is 20.

In this study we are interested in the dynamics of formation of groups. Another important feature of [KL13], 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 strictly increases 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 strictly increases their utility. If the group they join is empty then, 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 [KL13]. 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. Hence, any user cannot join another group. Furthermore, when $k \leq 3$, if $k$ users join an empty group then, their utility will be at most 2 and 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 in doing 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 total 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 [KL13].

## 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 [PS08] as a game-theoretic setting for studying the chromatic number in graphs. Specifically, the authors in [PS08] 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 [CKPS10, EGM12, KL13]. 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 [CNS18, FMZ17, Haj06, HJ17, MS17, OBI$^+$17]. We stress that while every coloring game has a Nash equilibrium that can be computed in polynomial-time [PS08], deciding whether a given Hedonic game admits a Nash equilibrium is NP-complete [Bal04]. In [DBHS06] the authors consider a special case of binary Hedonic games, close to our model, where every user considers another user either as a friend or as an enemy. However, unlike in our case, such preferences may be asymmetric (i.e., $u$ may consider $v$ as a friend, whereas $v$ considers $u$ as an enemy). This implies that a stable partition may not always exist, whereas it always does in our model. The authors in [DBHS06] have thus focused on the existence of core partitions (a.k.a., $k$-stable partitions for any $k$) under different utility functions. For other works about the existence of core coalitions in such games, and related ones with neutral preferences, see [OBI$^+$17] and the papers cited therein.

If the set of edges of the conflict graph is empty (edgeless conflict graph) then, there exists a unique $k$-stable partition, namely, that consisting of one group with all the users. In [KL13], it is proved that there always exists a $k$-stable partition for any conflict graph, but that it is NP-hard to compute one if $k$ is part of the input (this result was also proved independently in [EGM12]). Indeed, if $k$ is equal to the number of users then, a largest group in such a partition must be a maximum independent set of the conflict graph. In contrast, a $k$-stable partition can be computed in polynomial time for every fixed $k \leq 3$, by using simple *better-response dynamics* [PS08, EGM12, KL13]. 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 [KL13] 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 $G^\emptyset$ of order $n$. Prior to this work, no lower bound on $L(k, n)$ was known, and the analysis was limited to the use of a potential function that only applies when $k \leq 3$ [EGM12, KL13] 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 only depends on the size of the groups of the partitions generated. Following [Bry73], let an integer partition of $n \geq 1$, be a non-increasing sequence of integers $Q = (q_1, q_2, \ldots, q_n)$ such that $q_1 \geq q_2 \geq \ldots \geq q_n \geq 0$ and $\sum_{i=1}^n q_i = n$. If we rank the groups by non-increasing order of their size then, there is a natural relation between partitions in groups and integer partitions (the sizes of the groups $X_i$ corresponding to the integers $q_i$ of the partition of $n$). By 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 upper bound is certainly not tight. Indeed, we considered a worst-case scenario where all possible partitions could be attained with a single sequence of $k$-deviations.

| $k$ | Prior to our work | **Our results** | |
|---|---|---|---|
| 1 | $\mathcal{O}(n^2)$ [KL13] | exact analysis, that implies $L(1,n) \sim \frac{(2n)^{3/2}}{3}$ | Theorem 8 |
| 2 | $\mathcal{O}(n^2)$ [KL13] | exact analysis, that implies $L(2,n) \sim \frac{(2n)^{3/2}}{3}$ | Theorem 11 |
| 1-2 | $\mathcal{O}(n\alpha(G^-))$ [PS08] | $L(k,G^-) = \Omega(n\alpha(G^-))$ for some $G^-$ with $\alpha(G^-) = \mathcal{O}(\sqrt{n})$ | Theorem 14 |
| 3 | $\mathcal{O}(n^3)$ [EGM12, KL13] | $L(3,n) = \Omega(n^2)$ | Theorem 16 |
| $\geq 4$ | $\mathcal{O}(2^n)$ [KL13] | $L(k,n) = \Omega(n^{\Theta(\ln(n))})$, $L(k,n) = \mathcal{O}(\exp(\pi\sqrt{2n/3})/n)$ | Theorem 15 |

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

Nevertheless, 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 [KL13].

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 [GK86] in order 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 $\mathcal{O}(n^2)$.
  - In Section 4.2, we prove that any 2-deviation can be “replaced” (in some precise way) either by one or two 1-deviations. Therefore, $L(2,n) = L(1,n)$.
  - For $k = 1, 2$ and a general conflict graph $G^-$, the value of $L(k,G^-)$ depends on the independence number $\alpha(G^-)$ (cardinality of a largest independent set) of the conflict graph. In [PS08] it was proved that the convergence of the dynamics is in $\mathcal{O}(n\alpha(G^-))$. In the case of the edgeless conflict graph, we have seen that $L(1,n) = \mathcal{O}(n^{3/2})$. Hence, the preceding upper bound was not tight. Thus, we inferred that the convergence of the dynamics was in $\mathcal{O}(n\sqrt{\alpha(G^-)})$. Yet in fact we prove in Section 4.3 that, for any $\alpha(G^-) = \mathcal{O}(\sqrt{n})$, there exists a conflict graph $G^-$ with $n$ vertices and independence number $\alpha(G^-)$ for which we reach a stable partition after a sequence of at least $\Omega(n\alpha(G^-))$ 1-deviations.
- Finally, our main contribution is obtained for $k \geq 3$. Prior to our work, it was known that $L(3,n) = \mathcal{O}(n^3)$, that follows from another application of the potential function method [KL13]. But nothing proved that $L(3,n) > L(2,n)$, and in fact it was conjectured in [EGM12] that both values are equal. In Section 5, we prove (Theorem 16) that $L(3,n) = \Omega(n^2)$ and thus we show for the first time that deviations can delay convergence and that the gap between $k = 2$ and $k = 3$ obtained from the potential function method is indeed justified. It was also conjectured in [KL13] that $L(k,n)$ was polynomial in $n$ for $k$ fixed. In Section 5.1 we disprove this conjecture and prove in Theorem 15 that $L(4,n) = \Omega(n^{\Theta(\ln(n))})$. This shows that 4-deviations are responsible for a sudden complexity increase, as no polynomial bounds exist for $L(4,n)$.

# 2 Notations

**Conflict graph.** We refer to [BM08] for standard graph terminology. For the remainder of the paper, we assume that we are given a <u>conflict graph</u> $G^- = (V,E)$ where $V$ is the set of vertices (called users or players in the introduction) and edges represent the incompatibility relation (*i.e.*, an edge means that the two users are enemies). The number of vertices is denoted by $n = |V|$. The independence number of $G^-$, denoted $\alpha(G^-)$, is the maximum cardinality of an independent set in $G^-$. In particular, if $\alpha(G^-) = n$ then the conflict graph is edgeless, we denote it by $G^{\emptyset} = (V, E = \emptyset)$, and call it the empty graph.

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

We use in our proofs two alternative representations of the partition $P$. The partition vector associated with $P$ is defined as $\overrightarrow{\Lambda}(P) = (\lambda_n(P), \ldots, \lambda_1(P))$, where $\lambda_i(P)$ is the number of groups of size $i$. The integer partition associated with $P$ is defined as $Q = (q_1, q_2, \ldots, q_n)$, where $q_i = |X_i|$ and satisfies $q_1 \geq q_2 \geq \ldots \geq q_n \geq 0$ and $\sum_{i=1}^{n} q_i = n$.

In the example of Figure 1(a) we have a partition $P$ of the 12 vertices into 4 groups each of size 3. Therefore, $\lambda_3(P) = 4$ and $\lambda_i(P) = 0$ for $i \neq 3$; i.e. $\overrightarrow{\Lambda}(P) = (0, 0, 0, 0, 0, 0, 0, 0, 0, 4, 0, 0)$. The corresponding integer partition is $Q(P) = (3, 3, 3, 3, 0, 0, 0, 0, 0, 0, 0, 0)$. In the example of Figure 1(b) we have a partition $P'$ of the 12 vertices into one group of size 4 and 4 groups each of size 2. Hence, $\lambda_4(P') = 1$, $\lambda_2(P') = 4$ and $\lambda_i(P') = 0$ for $i \notin \{2, 4\}$; in other words, $\overrightarrow{\Lambda}(P') = (0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 4, 0)$. The corresponding integer partition is $Q(P') = (4, 2, 2, 2, 2, 0, 0, 0, 0, 0, 0, 0)$.

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

We next define a natural vector representation for $k$-deviations. The difference vector $\overrightarrow{\varphi}$ associated with a $k$-deviation $\varphi$ from $P$ to $P'$ is equal to $\overrightarrow{\varphi} = \overrightarrow{\Lambda}(P') - \overrightarrow{\Lambda}(P)$. In concluding this section, we define the $k$-deviations the most used in our proofs and explicit their difference vectors:

- $\alpha[p, q]$, the 1-deviation where a vertex leaves a group of size $q + 1$ for a group of size $p - 1$ (valid when $p \geq q + 2$). The non-zero coordinates of the associated vector $\overrightarrow{\alpha}[p, q]$ are: when $p > q + 2$, $\alpha_p = 1, \alpha_{p-1} = -1, \alpha_{q+1} = -1, \alpha_q = 1$; and for $p = q + 2$ (*i.e.*, $p - 1 = q + 1$), $\alpha_p = 1, \alpha_{p-1} = \alpha_{q+1} = -2, \alpha_q = 1$. For any $i \notin \{q, q + 1, p - 1, p\}$, $\alpha_i = 0$. (we omit for ease of reading the brackets $[p, q]$).

- $\gamma[p]$, the 3-deviation where one vertex in each of 3 groups of size $p - 1$ moves to a group of size $p - 3$ to form a new group of size $p$ (valid if there are at least 3 groups of size $p - 1$ and one of size $p - 3$). In that case the coordinates of $\overrightarrow{\gamma}[p]$ are: $\gamma_p = 1, \gamma_{p-1} = -3, \gamma_{p-2} = 3, \gamma_{p-3} = -1$, and $\gamma_i = 0$ for any $i \notin \{p - 3, p - 2, p - 1, p\}$.

- $\delta[p]$, the 4-deviation where one vertex in each of 4 groups of size $p - 1$ moves to a group of size $p - 4$ to form a new group of size $p$ (valid if there are at least 4 groups of size $p - 1$ and one of size $p - 4$). In that case the coordinates of $\overrightarrow{\delta}[p]$ are: $\delta_p = 1, \delta_{p-1} = -4, \delta_{p-2} = 4, \delta_{p-4} = -1$, and $\delta_i = 0$ for any $i \notin \{p - 4, p - 2, p - 1, p\}$. As an example, the move from the partition of Figure 1(a) to the partition of Figure 1(b), is a 4-deviation with difference vector $\overrightarrow{\delta}[4]$.

# 3 Preliminary results

In [KL13], the authors prove that there always exists a $k$-stable partition. We give another proof of this existence which will enable us to obtain a new upper bound. They also proved that it is NP-hard to compute a $k$-stable partition if $k$ is part of the input (this result was also proved independently in [EGM12]). In contrast, a $k$-stable partition can be computed in polynomial time for every fixed $k \leq 3$, by using simple *better-response dynamics* [PS08, EGM12, KL13]. The latter results question the role of the value of $k$ in the complexity of computing stable partitions.

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

$P_i$, we pick any one of these $k$-deviations $\varphi$ and we obtain a new partition $P_{i+1}$. If there is no $k$-deviation then, 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 $P_i$.

We now prove in Proposition 1 that better-response dynamics can be used for computing a $k$-stable partition *for every fixed* $k \geq 1$ (but not necessarily in polynomial time). It shows that for every fixed $k \geq 1$, the problem of computing a $k$-stable partition is in the complexity class PLS (Polynomial Local Search), that is conjectured to lie strictly between P and NP [JPY88]. Recall that the problem becomes NP-hard when $k$ is part of the input.

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

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

The proof of Algorithm 1 implies that $L(k, n)$ (the length of a longest sequence of $k$-deviations) is finite. We give here some elementary properties of $L(k, n)$.

**Property 2.** *$L(k, n)$ is always attained on the empty conflict graph $G^\emptyset$ of order $n$.*

*Proof.* It suffices to observe that any sequence of $k$-deviations on a conflict graph $G^-$ is also a sequence in the empty conflict graph with the same vertices. □

Note that the converse is not true as it can happen that some moves allowed in the empty conflict graph are not allowed in $G^-$ as they bring two enemies in the same group.

**Property 3.** *If $k' > k$, then $L(k', n) \geq L(k, n)$.*

*Proof.* We recall that by definition, a $k'$-deviation consists of a move of at most $k'$ vertices. In particular if $k' > k$ then, a $k$-deviation is also a $k'$-deviation. It implies that a sequence of $k$-deviations is also a sequence of $k'$-deviations. □

In fact we will see after that $L(1, n) = L(2, n) < L(3, n) \ll L(4, n)$. One can note that $k'$-deviations are more "powerful" than $k$-deviations in the sense that, a single $k'$-deviation might produce the same modification as multiple $k$-deviations. This does not affect the length of a longest sequence needed to reach a stable partition, but can reduce the length of a shortest sequence. For instance if we consider the empty graph $G^\emptyset$ of order $n$ and start with the partition consisting of $n$ groups of size 1 then, we need at least $(n-1)$ 1-deviations to reach the stable partition consisting of one group of size $n$, while we can reach the stable partition with only $\lceil (n-1)/k \rceil$ $k$-deviations.

Recall that we can associate with any partition $P = X_1, \ldots, X_i, \ldots, X_n$ of the vertices the integer partition $Q = (q_1, q_2, \ldots, q_n)$ such that $q_1 \geq q_2 \geq \ldots \geq q_n \geq 0$ and $\sum_{i=1}^{n} q_i = n$ by letting $q_i = |X_i|$.

The converse is not true in general; as an example it suffices to consider a partition with $q_1 > \alpha(G^-)$. However the converse is true when the conflict graph is empty; indeed it suffices to associate with an integer partition any partition of the vertices obtained by putting in the group $X_i$ a set of $q_i$ vertices .

Finally it is worth noting that partition vectors and integer vectors are equivalent representations of partitions as there is a one-to-one mapping between the partition vector and the integer partition associated with a partition.

We can now use the value $p_n$ of the number of integer partitions (see [HW79]) in order to obtain the following proposition:

**Proposition 4.** *$L(k, n) \leq p_n = \Theta((e^{\pi\sqrt{\frac{2n}{3}}})/n)$.*

*Proof.* By the proof of Proposition 1, in any sequence of $k$-deviations the partition vectors of the partitions obtained are all different. Furthermore as noted above, with two different partition vectors are associated two different integer partitions. Therefore, $L(k, n)$ is bounded by the number of integer partitions $p(n)$ which is $\Theta((e^{\pi\sqrt{\frac{2n}{3}}})/n)$ (see [HW79]). ☐

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 [KL13]. Note also that the number of integer partitions might be much larger than the maximum length of a sequence. Indeed, not all two consecutive (in the lexicographic ordering) partitions cannot appear in the same sequence. For instance for $n = 10$, if we have the partition consisting of two groups of size 5 then, we cannot reach with any $k$-deviation the partition with one group of size 6 and 4 groups of size 1. In Section 5.1 we give a lower bound $L(4, n) = \Omega(n^{\Theta(\ln(n))})$, but it still remains a large gap with the upper bound of Proposition 4.

# 4 Analysis for $k \leq 2$

In [KL13], the authors proved that for $k \leq 2$, Algorithm 1 converges to a stable partition in at most a quadratic time. They used the total utility (that is the sum of the utilities) as potential function. Indeed when performing a 1-deviation $\overrightarrow{\alpha}[p, q]$, 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. Thus, the total utility (potential function) increases by $2p - 2q - 2 \geq 2$ as $p \geq q + 2$.

Furthermore, in any partition, the utility of a vertex is at most $n - 1$ and the total utility is at most $n(n-1)/2$. Therefore the number of possible 1-deviations is bounded by $n(n-1)/4$, that is $L(1, n) = \mathcal{O}(n^2)$. Similarly one can also prove that when performing a 2-deviation the total utility increases by at least 2 (see also Claim 10). Hence, $L(2, n) = \mathcal{O}(n^2)$.

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 [Bry73], and is sometimes called the dominance ordering.

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

In [Bry73] 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 [OM16].

**Definition 5.** *(dominance ordering) Given two integer partitions of $n \geq 1$, $Q = (q_1, q_2, \ldots, q_n)$ and $Q' = (q'_1, q'_2, \ldots, q'_n)$, 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 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 6.** *Let $P$ be a partition of the vertices and $P'$ be the partition obtained after a 1-deviation $\varphi$. Then, the integer partition $Q' = Q(P')$ dominates $Q = Q(P)$.*

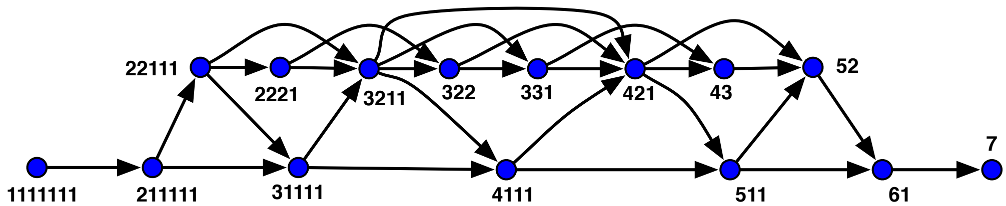


Figure 2: The lattice of integer partitions for $n = 7$, with the arcs associated with 1-deviations.

*Proof.* In the 1-deviation $\varphi$ a vertex $v$ moves from a group $X_k$ to a group $X_j$ with sizes $q_j = |X_j| \geq q_k = |X_k|$. W.l.o.g. we can assume 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 $X_k$ is the last group with size $|X_k|$. Thus, the integer partition $Q(P)$ associated with $P$ satisfies $q_1 \geq q_2 \ldots \geq q_{j-1} > q_j \geq q_{j+1} \geq \ldots \geq q_k > q_{k+1} \geq \ldots \geq q_n$. 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 with $P'$ has the same elements as $Q$ except $q'_j = q_j + 1$ and $q'_k = q_k - 1$. It implies that $Q'$ dominates $Q$. Note that this lemma holds for any conflict graph. □

In the example of Figure 2 (case $n = 7$), we have indicated the arcs between two integer partitions $Q(P)$ and $Q(P')$, where $P'$ is obtained from $P$ after some 1-deviation. We did not write in the figure the integers equal to 0. As an example, consider the partition $P$ with one group of size 3, one of size 2 and two of size 1. The integer partition associated with $P$ is $Q(P) = (3, 2, 1, 1, 0, 0, 0)$. Let $\varphi$ be the 1-deviation where a vertex in a group of size 1 moves to the other group of size 1. We obtain the partition $P'$ with one group of size 3 and two of size 2. The integer partition associated with $P'$ is $Q(P') = (3, 2, 2, 0, 0, 0, 0)$ which dominates $Q(P)$. We can also consider the other possible 1-deviations where a vertex in a group of size 1 (resp. 1, 2) moves to the group of size 2 (resp. 3, 3) getting respectively the integer partitions $(3, 3, 1, 0, 0, 0, 0)$ (resp. $(4, 2, 1, 0, 0, 0, 0)$, $(4, 1, 1, 1, 0, 0, 0)$).

For our next result we need to refine Def. 5, as follows. Let $Q, Q'$ be two integer partitions of $n \geq 1$. If $Q$ dominates $Q'$ and there is no other integer partition $Q''$ that simultaneously dominates $Q'$ and is dominated by $Q$ then, we say that $Q$ covers $Q'$. For instance in Fig. 2, the integer partition $(3, 1, 1, 1, 1)$ dominates $(1, 1, 1, 1, 1, 1, 1)$, but it does not cover this integer partition. Indeed, the integer partition $(2, 1, 1, 1, 1, 1)$ dominates $(1, 1, 1, 1, 1, 1, 1)$ and it is also dominated by $(3, 1, 1, 1, 1)$. However since no such integer partition exists "between" $(2, 1, 1, 1, 1, 1)$ and $(1, 1, 1, 1, 1, 1, 1)$, we have that $(2, 1, 1, 1, 1, 1)$ covers $(1, 1, 1, 1, 1, 1, 1)$. Brylawski proved a nice characterization of the covering relation, that we will use in our next proof.

**Lemma 7.** *Let $G^{\emptyset}$ 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 with $Q$, there exists another partition $P'$ associated with $Q'$ such that $P'$ is obtained from $P$ by doing a sequence of 1-deviations.*

*Proof.* As proved in [Bry73], we have that if $Q'$ dominates $Q$ then there is a finite sequence of integer partitions $Q^0, \ldots, Q^r, \ldots, Q^s$, 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 $q^{r+1}_{j_r} = q^r_{j_r} + 1$ and $q^{r+1}_{k_r} = q^r_{k_r} - 1$. In fact as proved in [Bry73], this is equivalent to have $Q^{r+1}$ that covers $Q^r$.

The proof is now by induction on $r$, starting from any partition $P^0 = P$ associated with $Q$. For $r > 0$, we consider the partition $P^r$ associated with $Q^r$. Recall that $Q^r$ and $Q^{r+1}$ only differ in the two groups $X_{j_r}$ and $X_{k_r}$. As $q^{r+1}_{j_r} = q^r_{j_r} + 1$ and $q^{r+1}_{k_r} = q^r_{k_r} - 1$, $P^{r+1}$ 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 $Q$. A possible sequence of integer partitions satisfying the conditions of the proof is $Q^0 = Q$, $Q^1 = (3, 2, 1, 1, 0, 0, 0)$, $Q^2 = (4, 1, 1, 1, 0, 0, 0)$, $Q^3 = Q' = (5, 1, 1, 0, 0, 0, 0)$. Partition $P^1$ is obtained from $P^0$ by moving a vertex in 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 Property 2, the length of a longest sequence of 1-deviations with an empty conflict graph is the same as the length of a longest chain in the dominance lattice of integer partitions. Since it has been proven in [GK86] 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 8.** *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, $L(1, n) = 2\binom{m+1}{3} + mr$.*

In the example for $n = 7$, we have $m = 3$, $r = 1$ and so, $L(1, 7) = 11$. Such a longest sequence is $(1, 1, 1, 1, 1, 1, 1)$, $(2, 1, 1, 1, 1, 1, 0)$, $(2, 2, 1, 1, 1, 0, 0)$, then the sequence in the upper part of Figure 2 until $(5, 2, 0, 0, 0, 0, 0)$, then $(6, 1, 0, 0, 0, 0, 0)$ and $(7, 0, 0, 0, 0, 0, 0)$.

We note that the proof in [GK86] 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 integer partition in the lexicographic ordering, like for $n = 7$. Unfortunately this is not true. Indeed let $n = 9$. After 7 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 integer partition (in the lexicographic ordering), we get the integer 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 [GK86] implies that the following simple construction works for any $n$.

**Proposition 9.** *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 total utility.*

*Proof.* We need to introduce the terminology of [GK86]. 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 to 1-deviations.

- A $V$-step is corresponding to a user leaving her group of size $p-1$ for another group of size $p-1$, thereby increasing her utility from $p - 2$ to $p - 1$. In other words, the deviation vector of such 1-deviation is $\overrightarrow{\alpha}[p, p-2]$ for some $p$.

- An $H$-step is corresponding to a user leaving her group $X_i$ of size $q + 1$ for another group $X_j$ of size $p-1 \geq q+1$, but only if there is no other group of size between $q+1$ and $p-1$ ; in particular, if groups are ordered by non-increasing size, this means that $j = i - 1$. 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 total utility has to increase by at least two, and this is attained if and only if the deviation vector is $\overrightarrow{\alpha}[p, p-2]$ for some $p$; equivalently, this move is corresponding to a V-step.

Furthermore, if no such a move is possible, then we claim that any 1-deviation $\overrightarrow{\alpha}[p, q]$ that minimizes the increase of the total utility is corresponding to an $H$-step. Indeed, we have $q+1 < p-1$ and, more generally, all the groups have pairwise different sizes (otherwise, a $V$-step would have been possible). Furthermore, if there were a group of size $s$, with $q + 1 < s < p - 1$, then this would contradict our assumption that the 1-deviation $\overrightarrow{\alpha}[p, q]$ minimizes the increase of the total utility.

As proved, *e.g.*, by Brylawski [Bry73], 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 (see also the proof of Lemma 7 for a different formulation).

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 [GK86]) it can be proved that as soon as no move $\overrightarrow{\alpha}[p, p-2]$ (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 $\overrightarrow{\alpha}[p, p-2]$ for some $p$, and there will be no other groups of size $p-1$ 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 [GK86]. Finally, the main result in [GK86] is that every $HV$-chain is of maximum length. □

In what follows, we will reuse part of the construction in [GK86] for proving Theorem 12.

## 4.2 Analysis for k = 2

Interestingly we will prove that any 2-deviation can be replaced either by one or two 1-deviations. This will allow us to prove in Theorem 11 that $L(2, n) = L(1, n)$.

**Claim 10.** *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, 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 $X_i$.
- Case 2: $|X_k| = |X_i| - 1 = |X_j| - 1 = p - 2$ and $X_i = X_j$. In that case, the effect of the 2-deviation is to replace the group $X_i$ of size $p - 1$ with a group of size $p - 3$ and to replace the group $X_k$ of size $p - 2$ with a group of size $p$. The difference vector $\overrightarrow{\varphi}$ associated with the 2-deviation has as non-null coordinates $\varphi_p = 1, \varphi_{p-1} = -1, \varphi_{p-2} = -1, \varphi_{p-3} = 1$. We obtain the same effect by doing the 1-deviation $\overrightarrow{\alpha}[p, p - 3]$ where a vertex leaves $X_k$ to join $X_i$.
- 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 2-deviation is to replace the 2 groups $X_i$ and $X_j$ of size $p - 1$ with two groups of size $p - 2$ and to replace the group $X_k$ of size $p - 2$ with a group of size $p$. The difference vector $\overrightarrow{\varphi}$ associated with the 2-deviation has as non-null coordinates $\varphi_p = 1, \varphi_{p-1} = -2, \varphi_{p-2} = 1$. We obtain the same effect by doing the 1-deviation $\overrightarrow{\alpha}[p, p - 2]$ where a vertex leaves $X_j$ to join $X_i$.

Note that the fact that $G^-$ is empty is needed for the proof. Indeed, in case 2 it might happen that all the vertices of $X_k$ have some enemy in $X_i$ and so, the 1-deviation we describe is not valid. Similarly, in case 3, it might happen that all the vertices of $X_i$ have some enemy in $X_j$ and so, the 1-deviation we describe is not valid. ⋄

**Theorem 11.** *$L(2, n) = L(1, n)$.*

*Proof.* Clearly, $L(2, n) \geq L(1, n)$ as any 1-deviation is also a 2-deviation. By Property 2, the value of $L(2, n)$ is obtained when the conflict graph $G^-$ is empty. In that case, Claim 10 implies that $L(2, n) \leq L(1, n)$. □

## 4.3 Analysis for k ≤ 2 and a general conflict graph

By using the potential function introduced at the beginning of this section, Panagopoulou and Spirakis ([PS08]) proved that for every conflict graph $G^-$ with independence number $\alpha(G^-)$, the convergence of the dynamics is in $\mathcal{O}(n\alpha(G^-))$. Indeed as we have seen each 1-deviation increases the total utility by at least 2. But the total utility of a stable partition is at most $n(\alpha(G^-) - 1)$ as the groups have maximum size $\alpha(G^-)$. If the conflict graph is empty, we have seen that $L(1, n) = \Theta(n^{3/2})$, that is, in that case $\mathcal{O}(n\sqrt{\alpha(G^-)})$. This led one of us ([Duc16b], page 131) to conjecture that in the case of 1-deviations the worst time of convergence of the dynamics is in $\mathcal{O}(n\sqrt{\alpha(G^-)})$. We disprove the conjecture by proving the following theorem:

**Theorem 12.** *For $n = \binom{m+1}{2}$, there exists a conflict graph $G^-$ with $\alpha(G^-) = m = \Theta(\sqrt{n})$ and a sequence of $\binom{m+1}{3}$ valid 1-deviations, that is, a sequence of $\Omega(n^{\frac{3}{2}}) = \Omega(n\alpha(G^-))$ 1-deviations.*

*Proof.* The reader can follow the proof on the example given after for $m = 4$ and on Figure 3. We will use part of the construction of Greene and Kleitman ([GK86]). Namely, they prove that, if $n = \binom{m+1}{2}$, there is a sequence of $\binom{m+1}{3}$ 1-deviations transforming the partition $P_1$ consisting of n groups each of size 1 (the coordinates of $\overrightarrow{\Lambda}(P_1)$ satisfy $\lambda_1 = n$) into the partition $P_m$ consisting of $m$ groups, one of each possible size $s$ for $1 \leq s \leq m$ (the coordinates of $\overrightarrow{\Lambda}(P_m)$ satisfy $\lambda_s = 1$ for $1 \leq s \leq m$). Furthermore they prove that the moves used are $V$-steps (see the proof of Proposition 9) which are nothing else than $\overrightarrow{\alpha}[p, p-2]$ for some $p$ (one vertex leaves a group of size $p-1$ to join a group of the same size $p-1$). One can note that in such a move the utility increases exactly by 2, and as the total utility of $P_m$ is $\sum_{i=1}^{m} i(i-1) = (m+1)m(m-1)/3$, the number of moves is exactly $(m+1)m(m-1)/6$.

The conflict graph of the counterexample consists of $m$ complete graphs $K^j, 1 \leq j \leq m$, where $K^j$ has exactly $j$ vertices. An independent set is therefore formed by taking at 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 $s$ will be $X_s[m] = \bigcup x_{m+1-s}^j$ with $m+1-s \leq j \leq m$. Thus, these groups are independent sets.

The idea of the proof is to create the groups of $P_m$ by moving in successive phases $p$, $1 \leq p \leq m-1$, the vertices $x_i^{p+1}$ starting from $x_1^{p+1}$ until $x_p^{p+1}$. More precisely, 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 $s$ for $1 \leq s \leq p$, namely $X_s[p] = \bigcup x_{p+1-s}^j$ with $p+1-s \leq j \leq p$, (these groups form a partition of the vertices $x_i^j$ with $1 \leq i \leq j \leq p$) plus $n - p(p+1)/2$ groups of size 1 (corresponding to the vertices $x_i^j$ with $1 \leq i \leq j$ and $p+1 \leq j \leq m$). We will show in Claim 13 that we can transform the partition $P_p$ into $P_{p+1}$ with $p(p+1)/2$ valid 1-deviations.

**Claim 13.** *For any $1 \leq p \leq m-1$, there exists a sequence of $p(p+1)/2$ valid 1-deviations which transform the partition $P_p$ into $P_{p+1}$.*

Proof. (See example given after the proof.) In a first subphase, we move the vertex $x_1^{p+1}$ successively to the group $X_1[p]$, then to $X_2[p]$ and so on until $X_p[p]$. After these $p$ 1-deviations we have created the group $X_{p+1}[p+1] = X_p[p] \cup \{x_1^{p+1}\} = \bigcup x_1^j$ for $1 \leq j \leq p+1$. In the general subphase $1 \leq h \leq p$, we move via $(p+1-h)$ 1-deviations the vertex $x_h^{p+1}$ successively to the group $X_1[p]$, until $X_{p+1-h}[p]$, therefore creating the group $X_{p+2-h}[p+1] = \bigcup x_h^j$ for $h \leq j \leq p+1$. To conclude the proof we have to note that all these 1-deviations are valid, as we move a vertex from $K^{p+1}$ successively to groups containing vertices from $\bigcup K^j$, for $1 \leq j \leq p$ (and there is no edge between this vertex and the vertices of these groups). ◇

Thus, by applying $m-1$ times the previous claim, we can transform the partition $P_1$ into $P_m$ with $\sum_{p=1}^{m-1} p(p+1)/2 = \binom{m+1}{3}$ 1-deviations, therefore proving the theorem. □

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

Partition $P_1$ consists of the group $X_1[1] = \{x_1^1\}$. Following the proof of Claim 13 for $p = 1$, the partition $P_2$ is obtained by moving $x_1^2$ to $X_1[1]$, thereby creating the group $X_2[2] = \{x_1^1, x_1^2\}$ (the group of size 1 of $P_2$ being $X_1[2] = \{x_2^2\}$). By the proof of Claim 13 for $p = 2$, the partition $P_3$ is obtained by doing the 3 following 1-deviations: first we move $x_1^3$ to $X_1[2]$ and then $X_2[2]$, thereby creating the group $X_3[3] = \{x_1^1, x_1^2, x_1^3\}$; then we move $x_2^3$ to $X_1[2]$ for creating the group $X_2[3] = \{x_2^2, x_2^3\}$ (the group of size 1 of $P_3$ being $X_1[3] = \{x_3^3\}$). Similarly, the partition $P_4$ is obtained by doing the following 6 1-deviations: first we move $x_1^4$ to $X_1[3]$, then $X_2[3]$ and $X_3[3]$ for creating the group $X_4[4] = \{x_1^1, x_1^2, x_1^3, x_1^4\}$; then we move $x_2^4$ to $X_1[3]$ and then $X_2[3]$ for creating the group $X_3[4] = \{x_2^2, x_2^3, x_2^4\}$, and finally we move $x_3^4$ to $X_1[3]$ for creating the group $X_2[4] = \{x_3^3, x_3^4\}$ (the group of size 1 of $P_4$ being $X_1[4] = \{x_4^4\}$).

We can prove a theorem analogous to Theorem 12 when $\alpha(G^-) = \mathcal{O}(\sqrt{n})$.

**Theorem 14.** *For any $\alpha = \mathcal{O}(\sqrt{n})$, there exists a conflict graph $G^-$ with $n$ vertices and independence number $\alpha(G^-) = \alpha$ and a sequence of at least $\Omega(n\alpha)$ 1-deviations to reach a stable partition.*

*Proof.* Let $G_0^-$ be the graph of Theorem 12 for $m = \alpha$. $G_0^-$ has $n_0 = \mathcal{O}(\alpha^2)$ vertices, independence number $\alpha$, and furthermore there exists a sequence of $\Theta(\alpha^3)$ valid 1-deviations for $G_0^-$. Let $G^-$ be the graph obtained by taking the complete join of $k = n/n_0$ copies of $G_0^-$ (*i.e.*, we add all possible edges between every two copies of $G_0^-$). By construction, $G^-$ has order $n = kn_0 = \mathcal{O}(k\alpha^2)$ and the same independence number $\alpha$ as $G_0^-$. Furthermore, there exists a sequence of $k\Theta(\alpha^3) = \Omega(n\alpha)$ valid 1-deviations for $G^-$. □

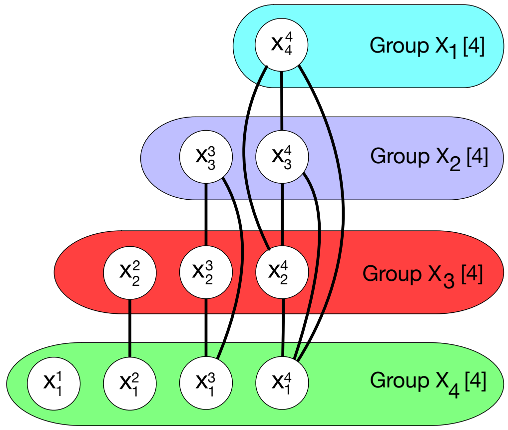


Figure 3: Illustration for Example 4.

Note that in any 2-deviation the total utility increases by at least 2. Hence, the number of 2 deviations is also at most $\mathcal{O}(n\alpha(G^-))$ when the conflict graph has independence number $\alpha(G^-)$. This bound is attained by using only 1-deviations as proved in Theorem 14, which is also valid for $k = 2$.

# 5 Lower bounds for $k > 2$

The classical dominance ordering does not suffice to describe all $k$-deviations as soon as $k \geq 3$. Indeed Lemma 6 is no more valid. Let $P$ be a partition of the vertices and $P'$ be the partition obtained after a $k$-deviation $\varphi$. Then, if $k \geq 3$ the integer partition $Q' = Q(P')$ does not necessarily dominate $Q = Q(P)$. Let us give an example. As noted before, there is only one $k$-stable partition $P_{max}$ in the empty conflict graph $G^{\emptyset}$, namely, the one consisting of one group of size $n$, with integer partition $Q_{max} = (n, 0, \ldots, 0)$ and partition vector $(1, 0, \ldots, 0)$. Let $d(Q)$ be the length of a longest sequence in the dominance lattice from the integer partition $Q$ to the integer partition $Q_{max}$. For $k = 4$, let $P$ be the partition consisting of 4 groups of size 4 and one group of size 1. We have $Q(P) = Q = (4, 4, 4, 4, 1)$ (we omit the 0's). Apply the 4-deviation $\delta[5]$ where one vertex of each group of size 4 joins the group of size 1; it leads to the partition $P'$ with integer partition $Q' = (5, 3, 3, 3, 3)$. $Q$ is uniquely covered in the dominance lattice by the integer partition $(5, 4, 4, 3, 1)$ while $Q'$ is at distance 3 from it via $(5, 4, 3, 3, 2)$ and $(5, 4, 4, 2, 2)$ and so, $d(Q') \geq d(Q) + 2$. Furthermore we note that the total utility of $P$ is 48 while that of $P'$ is 44, a smaller value.

Prior to our work, it was known that $L(3, n) = \mathcal{O}(n^3)$ [KL13]. But it was unknown whether $L(3, n) > L(2, n)$, and in fact it was conjectured in [EGM12] that both values are equal. It was also conjectured in [KL13] that $L(k, n)$ was polynomial in $n$ for $k$ fixed. We disprove this conjecture and prove in Theorem 15 that 4-deviations are responsible for a sudden complexity increase. Indeed, we prove that no polynomial bounds exist for $L(4, n)$ (Theorem 15).

**Theorem 15.** $L(4, n) = \Omega(n^{\Theta(\ln(n))})$.

The tools developed in the proof of Theorem 15 can also be used in the case $k = 3$ in order to prove for the first time that deviations can delay convergence and that the gap between $k = 2$ and $k = 3$ obtained from the potential function method is indeed justified (Theorem 16).

**Theorem 16.** $L(3, n) = \Omega(n^2)$.

The main idea of the proofs of these theorems is to do repeated ordered sequences (called cascades) of deviations similar to the one given in the example above (all these sequences are obtained by using a similar procedure, but with a first group of different size).

We give the proof of Theorem 16 in Section 5.2. In the following, we give the proof of Theorem 15 for $k = 4$. We use ordered sequences (cascades) of 4-deviations $\delta[p]$ (defined in the introduction) and various additional tricks such as the repetition of the process by using cascades of cascades. For completeness, we recall here the definition of $\delta[p]$.

**Definition of $\delta[p]$:** Consider a partition $P$ containing at least 4 groups of size $p-1$ and 1 group of size $p-4$. In the 4-deviation $\delta[p]$ one vertex in each of the 4 groups of size $p-1$ moves to the group of size $p-4$ to form a new group of size $p$.

The example given at the beginning of this section corresponds to the case $p = 5$ and the one given in Figure 1 to the case $p = 4$. Our motivation for using $\delta[p]$ as a basic building block for our construction is that it is the only type of 4-deviation which decreases the total utility.

Figure 4 gives a visual description of the cascades used in the proof. The proof uses a value $t$ which will be chosen at the end of the proof. In the figure, $t = 16$. Here we first do an ordered sequence of $t$ translated 4-deviations, with associated vectors $\overrightarrow{\delta}[p], \overrightarrow{\delta}[p-1], \ldots, \overrightarrow{\delta}[p-t+1]$ represented by black rectangles. The cascade obtained, called $\overrightarrow{\delta}^1[p,t]$, is represented by the red big rectangle. Then we do an ordered sequence of $(t-2)$ such cascades, i.e. $\overrightarrow{\delta}^1[p,t], \overrightarrow{\delta}^1[p,t-1], \ldots, \overrightarrow{\delta}^1[p,t-3]$, that are represented by red rectangles. In doing so we get the cascade $\overrightarrow{\delta}^2[p,t-2]$ which contains $16 \times 14 = 224$ 4-deviations. We apply some 1-deviations represented in white to get a deviation called $\overrightarrow{\zeta}^2[p]$, represented by the big yellow rectangle, which will have the so-called <u>Nice Property</u> (Definition 24) (few non-zero coordinates and symmetry properties) enabling us to do recursive constructions.

We repeat the process by doing ordered sequences of this new cascade $\overrightarrow{\zeta}^2$, but with translations of 2 instead of 1. More precisely we do successively the 7 cascades $\overrightarrow{\zeta}^2[p], \overrightarrow{\zeta}^2[p-2], \overrightarrow{\zeta}^2[p-4], \ldots, \overrightarrow{\zeta}^2[p-12]$ represented by yellow rectangles. After some additional 1-deviations (represented in white) we get the cascade called $\overrightarrow{\zeta}^3[p]$, that is represented by the big blue rectangle; this cascade still has the Nice property. Then, we do a cascade of 5 of these cascades, but now with a translation of 3, namely $\overrightarrow{\zeta}^3[p], \overrightarrow{\zeta}^3[p-3], \ldots, \overrightarrow{\zeta}^3[p-15]$ (represented by blue rectangles). After some additional 1-deviations (represented in white) we get the cascade called $\overrightarrow{\zeta}^4[p]$ (represented by the big green rectangle) that still has the Nice property. Finally the figure represent the ordered sequence of 3 of these cascades (with a translation of 5), namely $\overrightarrow{\zeta}^4[p], \overrightarrow{\zeta}^4[p-5], \overrightarrow{\zeta}^4[p-10]$, that are represented by green rectangles. After some 1-deviations (represented in white) we get the bigger cascade called $\overrightarrow{\zeta}^5[p]$ (represented by the turquoise rectangle). We continue the construction with a sequence of 3 of them (with a translation of 8), namely $\overrightarrow{\zeta}^5[p], \overrightarrow{\zeta}^5[p-8], \overrightarrow{\zeta}^5[p-16]$, plus some 1-deviations in order to get a cascade $\overrightarrow{\zeta}^6[p]$ (represented by the big grey rectangle). The reader has to realize that, in this example, $\overrightarrow{\zeta}^6[p]$ contains 3 cascades $\overrightarrow{\zeta}^5[p]$, each containing 3 cascades $\overrightarrow{\zeta}^4[p]$, each containing 5 cascades $\overrightarrow{\zeta}^3[p]$, each containing 7 cascades $\overrightarrow{\zeta}^2[p]$. Altogether the cascade $\overrightarrow{\zeta}^6[p]$ of this example contains $3 \times 3 \times 5 \times 7 \times 224 = 70560$ 4-deviations $\delta[p]$ plus many 1-deviations.

## 5.1 Case k = 4. Proof of Theorem 15

Note that after the 4-deviation $\delta[p]$ we get a new partition $P'$ with one more group of size $p$, 4 less groups of size $p-1$, 4 more groups of size $p-2$, and one less group of size $p-4$. This is expressed by the coordinates of the associated difference vector (where we omit the bracket $[p]$ for ease of reading).

**Difference vector $\overrightarrow{\delta}[p]$:** The difference vector $\overrightarrow{\delta}[p]$ has the following coordinates: $\delta_p = 1, \delta_{p-1} = -4, \delta_{p-2} = 4, \delta_{p-4} = -1$, and $\delta_j = 0$ for all other values of $j$. See Table 2.

| ... | $\delta_p$ | $\delta_{p-1}$ | $\delta_{p-2}$ | $\delta_{p-3}$ | $\delta_{p-4}$ | ... |
|---|---|---|---|---|---|---|
| ...0 | 1 | -4 | 4 | 0 | -1 | 0... |

Table 2: Difference vector of $\delta[p]$.

**The cascade $\overrightarrow{\delta}^1[p,t]$:** We first do a cascade consisting of an ordered sequence of $t$ translated 4-deviations $\delta[p], \delta[p-1], \ldots, \delta[p-t+1]$, for some parameter $t \geq 4$, that will be chosen later to give the maximum number of 4-deviations.

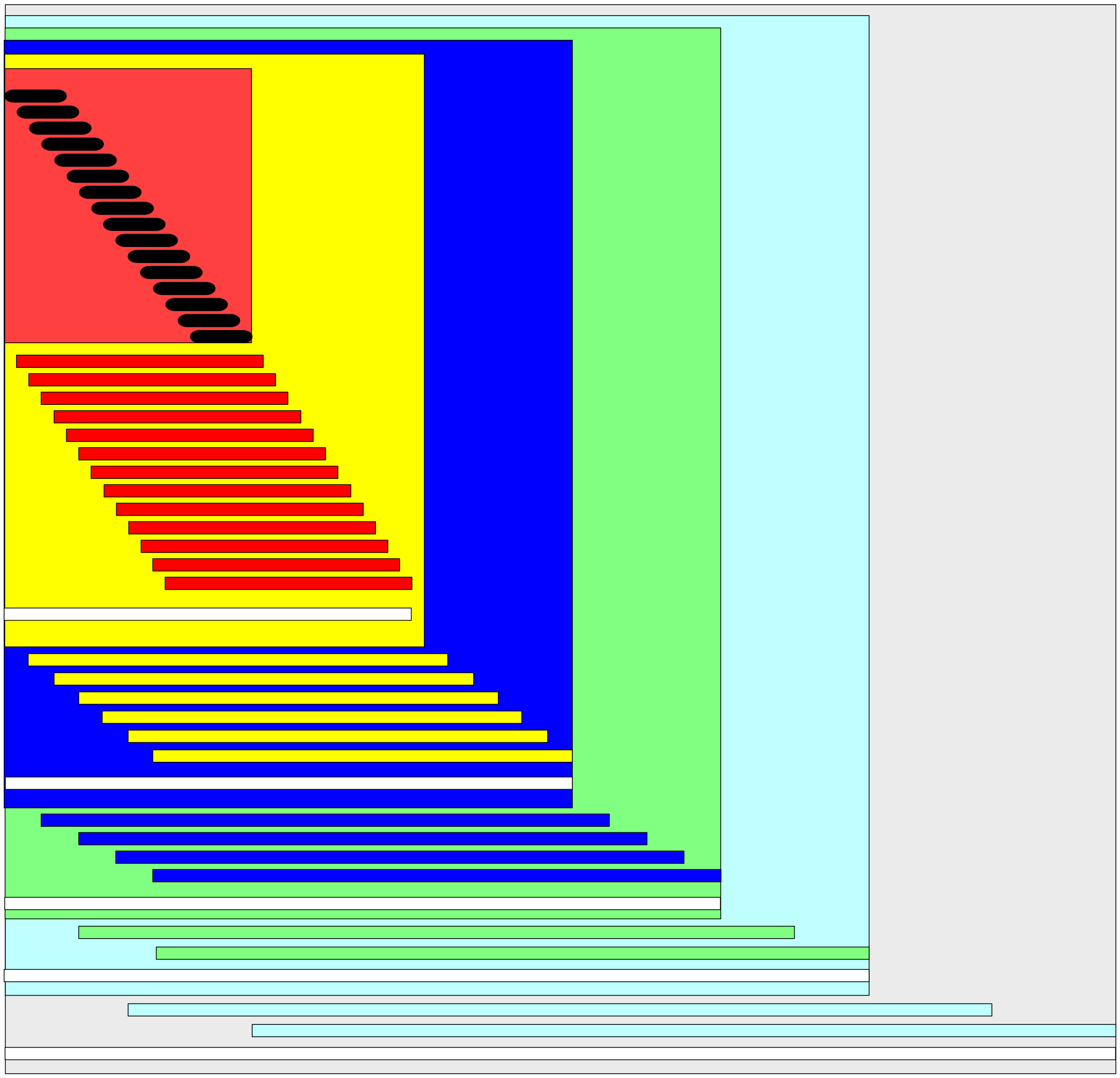

Figure 4: Cascades of cascades.

In what follows we represented a cascade $\Phi$ by the vector $\overrightarrow{\Phi}$ (as we are mainly interested in the sequence obtained) but we implicitly consider $\Phi$ as an ordered set of deviations which corresponding vector $\overrightarrow{\Phi}$ is equal to the sum of the corresponding vectors of the deviations.

The reader can follow the construction in Table 3 with $t = 7$. The coordinates of $\overrightarrow{\delta}^1[p,t]$ are given in Claim 17 and Table 4. We note that there are a lot of cancellations and only 8 non-zero coordinates. Indeed consider the groups of size $p-i$ for $4 \leq i \leq t-1$; we have deleted such a group when doing the 4-deviation $\overrightarrow{\delta}[p+4-i]$, then created 4 such groups with $\overrightarrow{\delta}[p+2-i]$, then deleted 4 such groups with $\overrightarrow{\delta}[p+1-i]$, and finally created one with $\overrightarrow{\delta}[p-i]$. The reader can follow these cancellations in Table 3 for $i = 4, 5, 6$. The variation of the number of groups of a given size $p-i$ (which corresponds to the coordinate $\delta^1_{p-i}$) is obtained by summing the coefficients appearing in the corresponding column and is 0 for $p-4, p-5, p-6$.

**Claim 17.** *For $4 \leq t \leq p-3$, the coordinates of the cascade $\overrightarrow{\delta}^1[p,t] = \sum_{i=0}^{t-1} \overrightarrow{\delta}[p-i]$ satisfy: $\delta^1_p = 1$, $\delta^1_{p-1} = -3$, $\delta^1_{p-2} = 1$, $\delta^1_{p-3} = 1$, $\delta^1_{p-t} = -1$, $\delta^1_{p-t-1} = 3$, $\delta^1_{p-t-2} = -1$, $\delta^1_{p-t-3} = -1$, and $\delta^1_j = 0$ for all the other values of $j$ (see Table 4).*

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

Table 3: Computation of $\delta^1[p,7]$.

| ... | $\delta^1_p$ | $\delta^1_{p-1}$ | $\delta^1_{p-2}$ | $\delta^1_{p-3}$ | ... | $\delta^1_{p-t}$ | $\delta^1_{p-t-1}$ | $\delta^1_{p-t-2}$ | $\delta^1_{p-t-3}$ | ... |
|---|---|---|---|---|---|---|---|---|---|---|
| ...0 | 1 | -3 | 1 | 1 | 0...0 | -1 | 3 | -1 | -1 | 0... |

Table 4: Difference vector $\delta^1[p,t]$.

Proof. We have $\delta^1_j = \sum_{i=0}^{t-1} \delta_j[p-i]$. For a given $j$, $\delta_j[p-i] = 0$ except for the following values of $i$ such that $0 \le i \le t-1$: $i = p-j$ where $\delta_j[j] = 1$; $i = p-j-1$ where $\delta_j[j+1] = -4$; $i = p-j-2$ where $\delta_j[j+2] = 4$; $i = p-j-4$ where $\delta_j[j+4] = -1$ (in the table it corresponds to the non-zero values in a column, whose number is at most 4). Therefore, for $j > p$: $\delta^1_j = 0$ ; $\delta^1_p = 1$; $\delta^1_{p-1} = -4+1 = -3$ (since $t \ge 2$); $\delta^1_{p-2} = 4-4+1 = 1$ (since $t \ge 3$); $\delta^1_{p-3} = 0+4-4+1 = 1$ (since $t \ge 4$); for $p-4 \ge j \ge p-t+1$, $\delta^1_{p-j} = -1+0+4-4+1 = 0$; $\delta^1_{p-t} = -1+0+4-4 = -1$ (since $t \ge 4$ and $t \le p$); $\delta^1_{p-t-1} = -1+0+4 = 3$ (since $t \ge 3$ and $t \le p-1$); $\delta^1_{p-t-2} = -1+0 = -1$ (since $t \ge 2$ and $t \le p-2$); $\delta^1_{p-t-3} = -1$ (since $t \ge 1$ and $t < p-3$) and, for $j \le p-t-4$, $\delta^1_j = 0$. ◇

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

**Definition 18.** *Let $h$ be a positive integer and let $\overrightarrow{\Phi} = \sum_{j=1}^{s} \overrightarrow{\varphi}^j$ be a cascade consisting of $s$ ordered $k$-deviations. We call this cascade $h$-balanced if, for any $1 \le r \le s$, the sum of the $r$ first vectors, namely, $\sum_{j=1}^{r} \overrightarrow{\varphi}^j$, has all its coordinates greater than or equal to $-h$.*

The interest in this notion lies in the following Property 19:

**Property 19.** *Let $p_{\max}$ be the largest index $j$ that satisfies $\overrightarrow{\Phi}_j \ne 0$. If we start from a partition with at least $h$ groups of each size $j$ such that $1 \le j < p_{\max}$, then an $h$-balanced sequence is valid.*

Note that the $p_{\max}$ coordinate is positive as it corresponds to a creation of a group and so we do not need the existence of groups of size $p_{\max}$. For example, the cascade $\overrightarrow{\delta}^1[p,t]$ described before is 4-balanced. Thus, if we have 4 groups of each size $j$, $1 \le j < p$, then the sequence is valid. We have seen above that in fact we do need so many groups for a size $\le p-2$; but the use of $h$-balanced sequences will simplify the computations and enable to get recursive properties.

Note that a sequence is itself composed of subsequences, and the following lemma will be useful to bound the value $h$ of a sequence.

**Lemma 20.** *Let $\overrightarrow{\Phi}^1$ be an $h_1$-balanced sequence and $\overrightarrow{\Phi}^2$ be an $h_2$-balanced sequence. Then, $\overrightarrow{\Phi}^1 + \overrightarrow{\Phi}^2$ is a (max $\{h_1, h_2 - min_i \Phi^1_i\}$)-balanced sequence.*

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

**The cascade $\overrightarrow{\delta}^2[p, t-2]$:** We do now the following sequence of $t-2$ cascades $\overrightarrow{\delta}^2[p, t-2] = \sum_{i=0}^{t-3} \overrightarrow{\delta}^1[p-i, t]$. Altogether we have a sequence of $t(t-2)$ 4-deviations. There are a lot of cancellations, and in fact, as shown in Claim 21, $\overrightarrow{\delta}^2[p, t-2]$ has only 10 non-zero coordinates. Table 5 describes an example of computation of $\overrightarrow{\delta}^2[p, t-2]$ with $t = 7$.

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

Table 5: Computation of $\delta^2[p, 5]$.

**Claim 21.** *For $5 \leq t \leq \frac{p}{2}$, the coordinates of the cascade $\overrightarrow{\delta}^2[p, t-2] = \sum_{i=0}^{t-3} \overrightarrow{\delta}^1[p-i, t]$ satisfy: $\delta^2_p = 1$, $\delta^2_{p-1} = -2$, $\delta^2_{p-2} = -1$, $\delta^2_{p-t+2} = -1$, $\delta^2_{p-t+1} = 2$, $\delta^2_{p-t-1} = 2$, $\delta^2_{p-t-2} = 1$, $\delta^2_{p-2t+2} = 1$, $\delta^2_{p-2t+1} = -2$, $\delta^2_{p-2t} = -1$, and $\delta^2_j = 0$ for all the other values of $j$ (see Table 6). Furthermore this cascade is 4-balanced.*

| ... | $\delta^2_p$ | $\delta^2_{p-1}$ | $\delta^2_{p-2}$ | ... | $\delta^2_{p-t+2}$ | $\delta^2_{p-t+1}$ | $\delta^2_{p-t}$ | $\delta^2_{p-t-1}$ | $\delta^2_{p-t-2}$ | ... | $\delta^2_{p-2t+2}$ | $\delta^2_{p-2t+1}$ | $\delta^2_{p-2t}$ | ... |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | -2 | -1 | 0 | -1 | 2 | 0 | 2 | 1 | 0 | 1 | -2 | -1 | 0 |

Table 6: Difference vector of $\delta^2[p, t-2]$.

Proof. We have $\delta^2_j = \sum_{i=0}^{t-3} \delta^1_j[p-i, t]$. Furthermore, by the hypothesis we have $4 \leq t \leq p-i-4$ for every $0 \leq i \leq t-3$, and so, we can apply Claim 17. By using the values of $\delta^1_j[p-i, t]$, we get that: for $j > p$, $\delta^2_j = 0$; $\delta^2_p = 1$; $\delta^2_{p-1} = -3+1 = -2$; $\delta^2_{p-2} = 1-3+1 = -1$ (since $t \geq 5$); for $p-3 \geq j \geq p-t+3$, $\delta^2_j = 1+1-3+1 = 0$; $\delta^2_{p-t+2} = 1+1-3 = -1$; $\delta^2_{p-t+1} = 1+1 = 2$; $\delta^2_{p-t} = -1+1 = 0$; $\delta^2_{p-t-1} = 3-1 = 2$; $\delta^2_{p-t-2} = -1+3-1 = 1$; for $p-t-3 \geq j \geq p-2t+3$, $\delta^2_j = -1-1+3-1 = 0$; $\delta^2_{p-2t+2} = -1-1+3 = 1$; $\delta^2_{p-2t+1} = -1-1 = -2$, $\delta^2_{p-2t} = -1$, and for $j < p-2t$, $\delta^2_j = 0$.

By using Lemma 20 we get that $\overrightarrow{\delta}^2[p, t-2]$ is 7-balanced; but a careful analysis shows that this sequence is in fact 4-balanced. Indeed, we will prove by induction that $\overrightarrow{\delta}^2[p, r] = \sum_{i=0}^{r-1} \overrightarrow{\delta}^1[p-i, t]$ is 4-balanced for any $r \leq t-2$. That is true for $r = 1$, as $\overrightarrow{\delta}^1[p, t]$ is 4-balanced. Suppose that it is true for $r$. We have to prove that any subsequence $\overrightarrow{\Phi}$ of $\overrightarrow{\delta}^2[p, r+1] = \overrightarrow{\delta}^2[p, r] + \overrightarrow{\delta}^1[p-r-1, t]$ has all its coordinates $\geq -4$. That is true by induction hypothesis if $\overrightarrow{\Phi}$ is a subsequence of $\overrightarrow{\delta}^2[p, r]$. Otherwise $\overrightarrow{\Phi}$ consists of $\overrightarrow{\delta}^2[p, r]$ plus a subsequence of $\overrightarrow{\delta}^1[p-r-1, t]$. All the coordinates of $\overrightarrow{\delta}^2[p, r]$ are at least $-3$, because the sum of the coordinates in any consecutive subsequence of $\overrightarrow{\delta}^1[p, t]$ is at least $-3$. Furthermore, as seen before the coordinates of $\overrightarrow{\delta}^1[p-r-1, t]$ are greater than $-1$, except, for $j = p-r-2$, where the coordinate is $\geq -4$. Therefore, for all values of $j \neq p-r-2$, $\Phi_j \geq -3-1 = -4$. For $j = p-r-2$, we have $\delta^2_{p-r-2}[p, r] = 1$ (case $r = 1$) or 2 (case $r > 1$) and so, $\Phi_{p-r-2} \geq 1-4 = -3$. In summary all the coordinates of any subsequence $\overrightarrow{\Phi}$ of $\overrightarrow{\delta}^2[p, r+1]$ are at least $-4$. Therefore by induction $\overrightarrow{\delta}^2[p, t-2]$ is 4-balanced. $\diamond$

At this stage we could continue and do a cascade of $\overrightarrow{\delta}^2[p, t-2]$, but there is no more the phenomenon of cancellation. In fact we will use the following "symmetrization" trick. We will transform the cascade

$\overrightarrow{\delta}^2[p, t-2]$ into a sequence $\overrightarrow{\zeta}^2[p]$ by doing some sequence of 1-deviations whose coordinates are given in Claim 22. 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 while preserving the Nice property.

For $p \geq q+2$, we recall that $\overrightarrow{\alpha}[p, q]$ is 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 $\overrightarrow{\alpha}^1[p, q, r] = \sum_{i=0}^{r-1} \overrightarrow{\alpha}[p-i, q+i]$ denote a cascade of $r$ such 1-deviations (we need $p-r+1 \geq q+r+2$). The coordinates of $\overrightarrow{\alpha}^1[p, q, r]$ are given in the following Claim 22.

**Claim 22.** *For $p-r \geq q+r+1$, $\overrightarrow{\alpha}^1[p, q, r] = \sum_{i=0}^{r-1} \overrightarrow{\alpha}[p-i, q+i]$ has only 4 non-zero coordinates, namely, $\alpha^1_p = 1$, $\alpha^1_{p-r} = -1$, $\alpha^1_{q+r} = -1$, and $\alpha^1_q = 1$.*

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

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

**Claim 23.** *For $5 \leq t \leq \frac{p-1}{2}$, the coordinates of the sequence $\overrightarrow{\zeta}^2[p] = \overrightarrow{\delta}^2[p, t-2] + \overrightarrow{\alpha}^1[p-1, p-2t-1, t-2] + \overrightarrow{\alpha}^1[p-1, p-2t, 2] + \overrightarrow{\alpha}^1[p-t+2, p-2t+1, 1] + \overrightarrow{\alpha}^1[p-t, p-t-3, 1]$ satisfy: $\zeta^2_p = 1$, $\zeta^2_{p-2} = -1$, $\zeta^2_{p-3} = -1$, $\zeta^2_{p-t} = 1$, $\zeta^2_{p-t-1} = 1$, $\zeta^2_{p-2t+2} = -1$, $\zeta^2_{p-2t+1} = -1$, $\zeta^2_{p-2t-1} = 1$ (see Table 8). Furthermore this cascade is still 4-balanced.*

| ... | $\zeta^2_p$ | $\zeta^2_{p-1}$ | $\zeta^2_{p-2}$ | $\zeta^2_{p-3}$ | ... | $\zeta^2_{p-t}$ | $\zeta^2_{p-t-1}$ | ... | $\zeta^2_{p-2t+2}$ | $\zeta^2_{p-2t+1}$ | $\zeta^2_{p-2t}$ | $\zeta^2_{p-2t-1}$ | ... |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ...0 | 1 | 0 | -1 | -1 | 0...0 | 1 | 1 | 0...0 | -1 | -1 | 0 | 1 | 0... |

Table 8: Difference vector $\zeta^2[p]$.

Proof. By Claim 22, we have the following coordinates:

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

Therefore, by using these values and the values of the coordinates of $\delta^2_j$ given in Claim 21, we get $\zeta^2_p = 1$, $\zeta^2_{p-1} = -2+1+1 = 0$, $\zeta^2_{p-2} = -1$, $\zeta^2_{p-3} = 0-1 = -1$, $\zeta^2_{p-t+2} = -1+1 = 0$, $\zeta^2_{p-t+1} = 2-1-1 = 0$, $\zeta^2_{p-t} = 0+1 = 1$, $\zeta^2_{p-t-1} = 2-1 = 1$, $\zeta^2_{p-t-2} = 1-1 = 0$, $\zeta^2_{p-t-3} = 0-1+1 = 0$, $\zeta^2_{p-2t+2} = 1-1-1 = -1$, $\zeta^2_{p-2t+1} = -2+1 = -1$, $\zeta^2_{p-2t} = -1+1 = 0$ $\zeta^2_{p-2t-1} = 0+1 = 1$.

To prove that $\overrightarrow{\zeta}^2[p]$ is 4-balanced, apply Lemma 20 with $\overrightarrow{\Phi}^1 = \overrightarrow{\delta}^2[p, t-2]$ and $\overrightarrow{\Phi}^2 = \overrightarrow{\alpha}^1[p-1, p-2t-1, t-2] + \overrightarrow{\alpha}^1[p-1, p-2t, 1] + \overrightarrow{\alpha}^1[p-t+2, p-2t+1, 1] + \overrightarrow{\alpha}^1[p-t, p-t-3, 1]$. We have that $h_1 = 4$ and furthermore all the coefficients of $\overrightarrow{\Phi}^1$ are greater than $-2$. Furthermore, we claim that $\overrightarrow{\Phi}^2$ is 2-balanced. Indeed, the four sequences of 1-deviations we use are clearly 1-balanced, and all their coordinates are at least $-1$. Therefore, the sum of any two such sequences is always 2-balanced. Now, we can observe that all the coordinates in $\overrightarrow{\alpha}^1[p-1, p-2t-1, t-2] + \overrightarrow{\alpha}^1[p-1, p-2t, 1]$ are at least $-1$. Therefore, we can apply Lemma 20 to $\overrightarrow{\alpha}^1[p-1, p-2t-1, t-2] + \overrightarrow{\alpha}^1[p-1, p-2t, 1]$ with $\overrightarrow{\alpha}^1[p-t, p-t-3, 1]$. Again, we can observe that all the coordinates in $\overrightarrow{\alpha}^1[p-1,p-2t-1,t-2]+\overrightarrow{\alpha}^1[p-1,p-2t,1]+\overrightarrow{\alpha}^1[p-t,p-t-3,1]$ are at least $-1$. Therefore, we can apply Lemma 20 to $\overrightarrow{\alpha}^1[p-1,p-2t-1,t-2]+\overrightarrow{\alpha}^1[p-1,p-2t,1]+\overrightarrow{\alpha}^1[p-t,p-t-3,1]$ with $\overrightarrow{\alpha}^1[p-t+2,p-2t+1,1]$.

Hence, $\overrightarrow{\zeta}^2[p]$ is $\max(4,2+2)=4$-balanced. ⋄

**Definition 24. Nice Property**: *Let $k \geq 2$ be a positive integer. We say that the sequence $\overrightarrow{\zeta}^k[p]$ has the Nice Property, if there exist three 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 all coordinates of $\overrightarrow{\zeta}^k$ are null except for:*

- $\zeta^k_p = \zeta^k_{p+1-2s(k)} = 1$,
- $\zeta^k_{p-a(k)} = \zeta^k_{p-b(k)} = \zeta^k_{p+1-2s(k)+b(k)} = \zeta^k_{p+1-2s(k)+a(k)} = -1$, *and*
- $\zeta^k_{p+1-s(k)} = \zeta^k_{p-s(k)} = 1$.

We note the symmetry of the coordinates, as for any $j$, $\zeta^k_{p-j} = \zeta^k_{p+1-2s(k)+j}$. As an example, the sequence $\overrightarrow{\zeta}^2[p]$ 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 $\overrightarrow{\zeta}^k[p]$ satisfying the Nice Property we can construct a sequence $\overrightarrow{\zeta}^{k+1}[p]$ having still the Nice Property.

**Claim 25. Main construction**: *Let $\overrightarrow{\zeta}^k[p]$ be a sequence satisfying the Nice Property with parameters $a(k), b(k), s(k)$ and with $p \geq 3s(k)-2-b(k)$. Then we can construct a sequence $\overrightarrow{\zeta}^{k+1}[p]$ satisfying the following properties:*

- $\overrightarrow{\zeta}^{k+1}[p]$ *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 largest even integer such that $r(k)a(k)+b(k) < s(k)-1$;*
- *if $\overrightarrow{\zeta}^k[p]$ is $h(k)$-balanced, then $\overrightarrow{\zeta}^{k+1}[p]$ is $(h(k)+1)$-balanced;*
- $\overrightarrow{\zeta}^{k+1}[p]$ *contains $r(k)+1$ sequences $\overrightarrow{\zeta}^k[p]$.*

Proof. (The reader can follow the proof on the example of Figure 4 that is detailed after the proof.) We will first do a cascade of $\overrightarrow{\zeta}^k[p]$, but we will take values of the parameters differing by a multiple of $a(k)$ in order for some of the coordinates to cancel (translation of $a(k)$) . Specifically, let us define $\overrightarrow{\Psi}^r = \sum_{j=0}^{r} \overrightarrow{\zeta}^k[p-ja(k)]$. We suppose for reasons explained below that $r$ is even and non-zero such that $p-b(k)-ra(k) > p+1-s(k)$. Using the values of Definition 24, we get the following values for the non-zero coordinates:

(1) $\psi^r_p = 1$; $\psi^r_{p-ja(k)} = -1+1 = 0$ for $0 < j \leq r$ (cancellation); $\psi^r_{p-(r+1)a(k)} = -1$;

(2) $\psi^r_{p+1-2s(k)+a(k)} = -1$; $\psi^r_{p+1-2s(k)-(j-1)a(k)} = 1-1 = 0$ for $0 < j \leq r$ (cancellation); $\psi^r_{p+1-2s(k)-ra(k)} = 1$;

(3) for $0 \leq j \leq r$, $\psi^r_{p-b(k)-ja(k)} = -1$;

(4) for $0 \leq j \leq r$, $\psi^r_{p+1-2s(k)+b(k)-ja(k)} = -1$;

(5) for $0 \leq j \leq r$, $\psi^r_{p+1-s(k)-ja(k)} = \psi^r_{p-s(k)-ja(k)} = 1$.

Since $a(k) < b(k) < 2a(k)$, all the indices of the coordinates are different provided we choose $r$ even and non-zero such that $p-b(k)-ra(k) > p+1-s(k)$ (that is equivalent to $ra(k)+b(k) < s(k)-1$). Furthermore we claim that all the indices are positive. Indeed on one hand the smallest index is $p+1-2s(k)-ra(k)$.

However on the other hand we have $p+1-2s(k)-ra(k) > p+2-3s(k)+b(k) \geq 0$ as we assume $p \geq 3s(k)-2-b(k)$.

Let us denote $a(k+1)=b(k)$; $b(k+1)=b(k)+a(k)$ and $s(k+1)=s(k)+a(k)r/2$.

Among the $4r+8$ non-zero coordinates of $\overrightarrow{\Psi}^r$, 8 of them correspond to the non-zero coordinates of a sequence having the Nice property with parameters $a(k+1), b(k+1)$ and $s(k+1)$, namely:

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

We claim that the $4r$ remaining non-zero values can be written as follows: for the values $-1$, in the form $\psi^r_{p-x_m}$ and $\psi^r_{p+1-2s(k+1)+x_m}$; and for the values 1, in the form $\psi^r_{p-y_m}$ and $\psi^r_{p+1-2s(k+1)+y_m}$ with $x_m < y_m$ $(0 \leq m \leq r-1)$. The proof of that assertion is technical and can be followed on the example given after the proof with $k=2$).

Indeed the $4r$ non-zero coordinates can be partitioned as follows: there are $r$ values equal to $-1$, namely $\psi^r_{p-b(k)-ja(k)} = -1$, for $2 \leq j \leq r$, and $\psi^r_{p-(r+1)a(k)} = -1$; then there are $2r$ values equal to 1, namely $\psi^r_{p+1-s(k)-ja(k)} = \psi^r_{p-s(k)-ja(k)} = 1$, for $0 \leq j \leq r, j \neq r/2$; finally there are $r$ values equal to $-1$, that are $\psi^r_{p+1-2s(k)+a(k)} = -1$ and $\psi^r_{p+1-2s(k)+b(k)-ja(k)} = -1$, for $0 \leq j \leq r-2$, respectively. We note that these values are disposed in a very symmetric way.

Let $x_m = b(k)+(m+2)a(k)$ for $0 \leq m \leq r-2$ and $x_{r-1} = (r+1)a(k)$. For $2 \leq j \leq r$, let $m=j-2$; then we get $p-b(k)-ja(k) = p-x_m$. We also have $p-(r+1)a(k) = p-x_{r-1}$. On the other side, for $0 \leq j \leq r-2$ let $m=r-j-2$; then we get $p+1-2s(k)+b(k)-ja(k) = p+1-2s(k)+b(k)-ra(k)+(m+2)a_k = p+1-2s(k+1)+x_m$ (recall that $2s(k+1)=2s(k)+ra(k)$). We also have $p+1-2s(k)+a(k) = p+1-2s(k+1)+(r+1)a(k) = p+1-2s(k+1)+x_{r-1}$.

The choice for the values $y_m$, $0 \leq m \leq r-1$ depends on the parity of $m$. Let $y_{2m'} = s(k)-1+m'a(k)$ and $y_{2m'+1} = s(k)+m'a(k)$ where $0 \leq m' \leq r/2-1$. For $0 \leq j \leq r/2-1$, let $m'=j$; then we get $p+1-s(k)-ja(k) = p-(s(k)-1+m'a(k)) = p-y_{2m'}$ and $p-s(k)-ja(k) = p-(s(k)+m'a(k)) = p-y_{2m'+1}$. In the same way for $r/2+1 \leq j \leq r$, let $m'=r-j$; then we get $p+1-s(k)-ja(k) = p+1-s(k)-ra(k)+m'a(k)) = p+1-2s(k+1)+s(k)+m'a(k)) = p+1-2s(k+1)+y_{2m'+1}$ and $p-s(k)-ja(k) = p-s(k)-ra(k)+m'a(k)) = p+1-2s(k+1)+s(k)-1+m'a(k)) = p+1-2s(k+1)+y_{2m'}$.

Furthermore, these $r$ quadruples of values can be canceled by adding to $\overrightarrow{\Psi}^r$ the $r$ sequences $\overrightarrow{\alpha}^1[p-x_m, p+1-2s(k+1)+x_m, y_m-x_m]$.

We claim that the resulting sequence with partition vector $\overrightarrow{\Psi}^r + \sum_{m=0}^{r-1} \overrightarrow{\alpha}^1[p-x_m, p+1-2s(k+1)+x_m, y_m-x_m]$ satisfies the Nice Property with parameters $a(k+1), b(k+1)$ and $s(k+1)$. Indeed the non-zero coordinates are exactly the 8 listed before. Furthermore, $a(k+1)=b(k) < b(k)+a(k) = b(k+1)$, $b(k+1) = b(k)+a(k) < b(k)+b(k) < 2a(k+1)$ and $b(k+1) = b(k)+a(k) < s(k)-1+a(k) \leq s(k+1)-1$ as $r \geq 2$. Finally, since we choose $p \geq 3s(k)-2-b(k)$, we get $p \geq (2s(k)+ra(k))+(s(k)-2-b(k)-ra(k)) > 2s(k+1)$ as $ra(k)+b(k) < s(k)-1$ and $p \geq 2s(k+1)-1$.

We now prove that $\overrightarrow{\zeta}^{k+1}[p]$ is $(h(k)+1)$-balanced. We first prove by induction that $\overrightarrow{\Psi}^r$ is $(h(k)+1)$-balanced. That is true for $r=0$ as $\overrightarrow{\zeta}^k[p]$ is $h(k)$-balanced. Then suppose it is true for some $r$; we apply Lemma 20 with $\overrightarrow{\Phi}^1 = \overrightarrow{\Psi}^r$ and $\overrightarrow{\Phi}^2 = \overrightarrow{\zeta}^k[p-(r+1)a(k)]$. We have that $h_1 = h(k)+1$ by our induction hypothesis and furthermore all the coefficients of $\overrightarrow{\Phi}^1$ are greater than $-1$; furthermore $\overrightarrow{\Phi}^2$ is $h(k)$-balanced and so, $\overrightarrow{\Psi}^{r+1}$ is $(\max(h(k)+1, h(k)+1) = h(k)+1)$-balanced. We add the 1-deviation $\overrightarrow{\alpha}^1[p-x_m, p+1-2s(k+1)+x_m, y_m-x_m]$ which is 1-balanced. Since $\max(h(k)+1, 1+1) = h(k)+1$, we still get a $(h(k)+1)$-balanced sequence.

Finally, by construction, we get that $\overrightarrow{\zeta}^{k+1}[p]$ contains $r(k)+1$ sequences $\overrightarrow{\zeta}^k[p]$. $\diamond$

**Example with t = 16.** (See Figure 4.) We start with $\overrightarrow{\zeta}^2[p]$, which satisfies the Nice property with $a(2) = 2, b(2) = 3$ and $s(2) = 17$, and whose coordinates are $\zeta^2_p = \zeta^2_{p-33} = 1$, $\zeta^k_{p-2} = \zeta^k_{p-3} = \zeta^k_{p-30} = \zeta^k_{p-31} = -1$, and $\zeta^k_{p-16} = \zeta^k_{p-17} = 1$. $r(2)$ is the largest even integer such that $2r(2) + 3 < 16$, that is $r(2) = 6$.

Thus, we can do a cascade of 7 sequences $\overrightarrow{\zeta}^2[p]$, translated by 2, and in doing so we get $\overrightarrow{\Psi}^6 = \sum_{j=0}^{6} \overrightarrow{\zeta}^2[p-2j]$. We have $\psi^6_p = \psi^6_{p-45} = 1$, $\psi^6_{p-3} = \psi^6_{p-5} = \psi^6_{p-40} = \psi^6_{p-42} = -1$, and $\psi^6_{p-22} = \psi^6_{p-23} = 1$. We also have 12 values $-1$ namely: $\psi^6_{p-7} = \psi^6_{p-9} = \psi^6_{p-11} = \psi^6_{p-13} = \psi^6_{p-15} = \psi^6_{p-14}$ and $\psi^6_{p-31} = \psi^6_{p-30} = \psi^6_{p-32} = \psi^6_{p-34} = \psi^6_{p-36} = \psi^6_{p-38}$, which can be written on the form $p - x_m$ and $p - 45 + x_m$ with $x_0 = 7, x_1 = 9, x_2 = 11, x_3 = 13, x_4 = 15, x_5 = 14$. We have 12 values 1 namely: $\psi^6_{p-16} = \psi^6_{p-17} = \psi^6_{p-18} = \psi^6_{p-19} = \psi^6_{p-20} = \psi^6_{p-21}$ and $\psi^6_{p-24} = \psi^6_{p-25} = \psi^6_{p-26} = \psi^6_{p-27} = \psi^6_{p-28} = \psi^6_{p-29}$, which can be written on the form $p - y_m$ and $p - 45 + y_m$ with $y_0 = 16, y_1 = 17, y_2 = 18, y_3 = 19, y_4 = 20, y_5 = 21$. Finally, the blue cascade $\overrightarrow{\zeta}^3[p] = \overrightarrow{\Psi}^6 + \sum_{m=0}^{5} \overrightarrow{\alpha}^1[p - x_m, p + 1 - 2s(k+1) + x_m, y_m - x_m]$ satisfies the Nice property with $a(3) = 3, b(3) = 5$ and $s(3) = 23$.

Then $r(3)$, the largest even integer such that $3r(2) + 5 < 22$, satisfies $r(3) = 4$. Applying the claim we get the green cascade $\overrightarrow{\zeta}^4[p]$, which satisfies the Nice property with $a(4) = 5, b(4) = 8$ and $s(4) = 29$. Then $r(4)$, the largest even integer such that $5r(4) + 8 < 28$, satisfies $r(4) = 2$. We apply once again the claim and in doing so we get the grey cascade $\overrightarrow{\zeta}^5[p]$, which satisfies the Nice property with $a(5) = 8, b(5) = 13$ and $s(5) = 34$. We can apply the claim again as $r(5)$, the largest even integer such that $8r(5) + 13 < 33$, satisfies $r(5) = 2$. Doing so, we get a cascade $\overrightarrow{\zeta}^6[p]$ which satisfies the Nice property with $a(6) = 13, b(6) = 21$ and $s(6) = 42$.

*Proof.* (**End of the proof of Theorem 15**) At this stage we have built a sequence $\overrightarrow{\zeta}^2[p]$ which satisfies the Nice Property with $a(2) = 2$, $b(2) = 3$ and $s(2) = t + 1$. Since $h(2) = 4$, it is 4-balanced. Furthermore, it contains $t(t-2)$ 4-deviations (see Claim 23). Then, for some well-chosen $K$ (to be defined later) we can apply $K - 2$ times the main construction (Claim 25) to construct a sequence $\overrightarrow{\zeta}^K[p]$ which satisfies the Nice Property with parameters $a(K)$, $b(K)$ and $s(K)$ and is $h(K)$-balanced.

On one hand since we want to upper bound the number of vertices, we need to upper bound $a(K)$, $b(K)$ and $s(K)$. We have $a(K) = b(K-1)$, $b(K) = b(K-1) + a(K-1) = b(K-1) + b(K-2)$ and so, we recognize the Fibonacci recurrence relation. The $jth$ Fibonacci number $F(j)$ satisfies:

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

Then, as $a(2) = 2 = F(3)$ and $b(2) = 3 = F(4)$, we get $a(K) = F(K+1)$ and $b(K) = F(K+2)$. In fact in what follows we will only use that $a(K) \leq 2^{K-1}$ and $b(K) \leq 2^K$. 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, $s(K+1) < (3/2) \times s(K)$. Therefore, we prove by induction that $s(K) < s(2)(3/2)^{K-2} = (t+1)(3/2)^{K-2}$.

Recall that we should have $p \geq 2s(K) - 1$. Hence, we choose $p = 2s(K)$. Furthermore by induction we have that $h(K) = K + 2$. Thus, by Proposition 19, it suffices to start with a partition containing at least $K + 2$ groups of each size $i$, $1 \leq i \leq p$. It is easy to obtain such a starting partition from the initial partition — which consists of $n$ groups of size 1 — by doing a sequence of 1-deviations of size $(K+2)p(p-1)/2$; indeed we can create a group of any size $i$ with $(i-1)$ 1-deviations. Therefore, we will take $n = (K+2)p(p-1)/2 = (K+2)s(K)(2s(K)-1)$. By using the inequality $s(K) < (t+1)(3/2)^{K-2}$ we get that:

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

On the other hand, we have to lower bound the number of deviations. By construction we have, for any value $k$, $\overrightarrow{\zeta}^{k+1}[p]$ contains $r(k) + 1$ sequences $\overrightarrow{\zeta}^k[p]$. Furthermore, $\overrightarrow{\zeta}^2[p]$ contains $t(t-2)$ 4-deviations. As a result, $\overrightarrow{\zeta}^K[p]$ contains $t(t-2)\prod_{k=2}^{K-1}(r(k)+1)$ 4-deviations.

We also recall that, for any $k \leq K$, $r(k)$ is the largest even integer $r$ such that $ra(k) + b(k) < s(k) - 1$. In particular, $r(k) \geq \lfloor \frac{s(k)-1-b(k)}{a(k)} \rfloor - 1$. By using the facts that $b(k) + 1 \leq 2a(k)$, $s(k) > s(2) - 1 = t$, and

$a(k) \leq a(K) < 2^{K-1}$, we get $r(k) \geq \frac{t}{2^{K-1}} - 3$. Then $\prod_{k=2}^{K-1}(r(k)+1) \geq (\frac{t}{2^{K-1}} - 2)^{K-2}$ and the number $D$ of deviations satisfies:

$$D = \Omega\left(t^2(\frac{t}{2^{K-1}} - 2)^{K-2}\right). \tag{2}$$

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

On one hand, we get by Equation 1 that $n = \mathcal{O}(t^2 \log_2(t)(3/2)^{\log_2(t)})$. Let $f(t) = t^2 \log_2(t)(3/2)^{\log_2(t)}$; then $\log_2(f(t)) = 2\log_2(t) + \log_2(\log_2(t)) + \log_2(t)(\log_2(3) - \log_2(2))$. By using that $\log_2(3) - \log_2(2) < 0.585$ and the fact that for $t$ large enough $\log_2(\log_2(t)) < 0.014 \log_2(t)$, we get for $t$ large enough $\log_2(f(t)) < 2.6\log_2(t))$ or equivalently $f(t) < t^{2.6}$. As a result, $n = \mathcal{O}(t^{2.6})$.

On the other hand we get by Equation 2: $D = \Omega((t^{1/2})^{\log_2(t)/2}) = \Omega(t^{\log_2(t)/4})$. Therefore, $D = \Omega(n^{c log_2(n)})$ with $c = \frac{1}{4\times(2.6)^2} \simeq 1/27$, thereby proving Theorem 15. $\square$

## 5.2 Case k = 3. Proof of Theorem 16

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). For completeness, we recall here the definition of $\gamma[p]$.

**Definition of** $\gamma[p]$**:** Consider a partition $P$ containing at least 3 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 $p$.

Let $P$ be the partition consisting of 3 groups of size 3 and one group of size 1. We have $Q(P) = Q = (3,3,3,1)$ (we omit the 0s). Apply the 3-deviation $\gamma[4]$ where one vertex of each group of size 3 joins the group of size 1; it leads to the partition $P'$ with integer partition $Q' = (4,2,2,2)$. Note that $Q$ and $Q'$ have the same total utility 18.

Thus, 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** $\overrightarrow{\gamma}[p]$**:** The difference vector $\overrightarrow{\gamma}[p]$ 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 9.

| ... | $\gamma_p$ | $\gamma_{p-1}$ | $\gamma_{p-2}$ | $\gamma_{p-3}$ | ... |
|---|---|---|---|---|---|
| ...0 | 1 | -3 | 3 | 1 | 0... |

Table 9: Difference vector of $\gamma[p]$.

**The cascade** $\overrightarrow{\gamma}^1[p,t]$**:** Like for $k = 4$, we do a cascade consisting of $t$ ordered 3-deviations $\gamma[p], \gamma[p-1], \gamma[p-2], \ldots, \gamma[p-t+1]$. We will denote this cascade by its difference vector $\overrightarrow{\gamma}^1[p,t] = \sum_{i=0}^{t-1} \overrightarrow{\gamma}[p-i]$.

The reader can follow the construction in Table 10 with $t = 5$. The coordinates of $\overrightarrow{\gamma}^1[p,t]$ are given in Claim 26 and Table 11. We note that there are a lot of cancellations and only 6 non-zero coordinates.

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

Table 10: Computation of $\gamma^1[p,5]$.

**Claim 26.** *For $3 \leq t \leq p-3$, the coordinates of the cascade $\overrightarrow{\gamma}^1[p,t] = \sum_{i=0}^{t-1} \overrightarrow{\gamma}[p-i]$ 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 all the other values of $j$. Furthermore, this cascade is 3-balanced (see Table 11).*

| ... | $\gamma_p^1$ | $\gamma_{p-1}^1$ | $\gamma_{p-2}^1$ | ... | $\gamma_{p-t}^1$ | $\gamma_{p-t-1}^1$ | $\gamma_{p-t-2}^1$ | ... |
|---|---|---|---|---|---|---|---|---|
| ...0 | 1 | -2 | 1 | 0...0 | -1 | 2 | -1 | 0... |

Table 11: Difference vector $\gamma^1[p,t]$.

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 values of $i$ such that $0 \leq i \leq t-1$: $i = p-j$ where $\gamma_j[j] = 1$; $i = p-j-1$ where $\gamma_j[j+1] = -3$; $i = p-j-2$ where $\gamma_j[j+2] = 3$; $i = p-j-3$ where $\gamma_j[j+3] = -1$ (in the table it corresponds to the consecutive non-zero values in a column which are at most 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 $p-3 \geq j \geq p-t+1$, $\gamma_{p-j}^1 = -1+3-3+1 = 0$; $\gamma_{p-t}^1 = -1+3-3 = -1$; $\gamma_{p-t-1}^1 = -1+3 = 2$; $\gamma_{p-t-2}^1 = -1$ and, for $j < p-t-2$, $\gamma_j^1 = 0$. Finally we note that the coordinates of any subsequence of the cascade, *i.e.* the coordinates of $\overrightarrow{\gamma}^1[p,r]$, are all at least $-1$ except $\gamma_{p-1}^1$: which is $-3$ when $r = 1$ and then $-2$, thereby proving that the cascade is 3-balanced. ◇

In our example we can do the cascade until $t = p-3$. However, it is better to choose $t$ smaller (we will see below that a good value for $t$ should satisfy $p - 4t + 3 \geq 0$) and repeat the previous cascade but translated (i.e., we do a cascade of cascades). More precisely, we now do the following sequence of $t-1$ cascades $\overrightarrow{\gamma}^2[p,t-1] = \sum_{i=0}^{t-2} \overrightarrow{\gamma}^1[p-i,t]$. Altogether we have a sequence of $t(t-1)$ 3-deviations.

In the example (see Table 12), we choose $t = 5$ and after having done $\overrightarrow{\gamma}^1[p,5]$ we do $\overrightarrow{\gamma}^1[p-1,5]$, $\overrightarrow{\gamma}^1[p-2,5]$ and $\overrightarrow{\gamma}^1[p-3,5]$. We can see again a phenomenon of cancelation as the coordinate $\gamma_{p-2}^2$ of this cascade of cascades equals 0. Similarly the coordinates $\gamma_{p-3}^2$ $\gamma_{p-7}^2$ and $\gamma_{p-8}^2$ of this cascade of cascades equal 0. This cancelation stays for all the other deviations. In the general case there are a lot of cancelations and in fact, as shown in the next Claim 27, $\overrightarrow{\gamma}^2[p,t-1]$ has only 6 non-zero coordinates.

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

Table 12: Example of cascade of cascades.

**Claim 27.** *For $3 \leq t \leq \frac{p-1}{2}$, the coordinates of the cascade $\overrightarrow{\gamma}^2[p,t-1] = \sum_{i=0}^{t-2} \overrightarrow{\gamma}^1[p-i,t]$ 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 other values of $j$. Furthermore this cascade is 3-balanced.*

Proof. We have $\gamma_j^2 = \sum_{i=0}^{t-2} \gamma_j^1[p-i,t]$. By using the values of $\gamma_j^1[p-i,t]$ given in Claim 26, 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 $j < p-2t$, $\gamma_j^2 = 0$.

| ⋯ | $\gamma^2_p$ | $\gamma^2_{p-1}$ | ⋯ | $\gamma^2_{p-t+1}$ | | $\gamma^2_{p-t-1}$ | ⋯ | $\gamma^2_{p-2t+1}$ | $\gamma^2_{p-2t}$ | ⋯ |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | -1 | 0 | -1 | 0 | 1 | 0 | 1 | -1 | 0 |

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

Here again we can see that after any number $r$ of 3-deviations the coordinates of the sequence are always at least $-3$. Indeed a subsequence either consists of a cascade $\overrightarrow{\gamma}^2[p,r] = \sum_{i=0}^{r-2} \overrightarrow{\gamma}^1[p-i,t]$ for some $r, 1 \leq r \leq t-1$, which has coordinates all greater than $-2]$; or of a cascade $\overrightarrow{\gamma}^2[p,r] = \sum_{i=0}^{r-2} \overrightarrow{\gamma}^1[p-i,t]$ for some $r, 1 \leq r \leq t-1$, plus a subsequence $\sum_{j=0}^{r'} \overrightarrow{\gamma}[p-r-j]$. The coordinates of the last subsequence are greater than $-1$ except the $p-r-1$ coordinate equal to $-3$ if $r'=0$ and $-2$ otherwise. But the $p-r-1$ coordinate of $\overrightarrow{\gamma}^2[p,r]$ is 1 and all the coordinates of the considered subsequence are greater than $-3$. ⋄

We can now repeat the cascade of translated cascades $\overrightarrow{\gamma}^2[p,t-1]$. More precisely, we now do the following sequence of $t-2$ cascades $\overrightarrow{\gamma}^3[p,t-2] = \sum_{i=0}^{t-3} \overrightarrow{\gamma}^2[p-i,t-1]$. Altogether we have a sequence of $t(t-1)(t-2)$ 3-deviations. In the example (see Table 12), after having done $\overrightarrow{\gamma}^2[p,4]$ we do $\overrightarrow{\gamma}^2[p-1,4]$ and $\overrightarrow{\gamma}^2[p-2,4]$. We can again see a phenomenon of cancelation as the coordinate $\gamma^3_{p-1}$ of this cascade of cascades equals 0. Similarly the coordinates $\gamma^3_{p-2}$,$\gamma^3_{p-10}$ and $\gamma^3_{p-11}$ of this cascade of cascades equal 0. In the general case there are a lot of cancelations. In fact, as shown in the next Claim 28, $\overrightarrow{\gamma}^3[p,t-2]$ has only 8 non-zero coordinates.

**Claim 28.** *For $3 \leq t \leq \frac{p+2}{3}$, the coordinates of the cascade $\overrightarrow{\gamma}^3[p,t-2] = \sum_{i=0}^{t-3} \overrightarrow{\gamma}^2[p-i,t-1]$ satisfy: $\gamma^3_p = 1$, $\gamma^3_{p-t+2} = \gamma^3_{p-t+1} = \gamma^3_{p-t} = -1$, $\gamma^3_{p-2t+3} = \gamma^3_{p-2t+2} = \gamma^3_{p-2t+1} = 1$, $\gamma^3_{p-3t+3} = -1$, and $\gamma^3_j = 0$ for all the other values of $j$. Furthermore this cascade is 4-balanced.*

| ⋯ | $\gamma^3_p$ | ⋯ | $\gamma^3_{p-t+2}$ | $\gamma^3_{p-t+1}$ | $\gamma^3_{p-t}$ | ⋯ | $\gamma^2_{p-2t+3}$ | $\gamma^2_{p-2t+2}$ | $\gamma^2_{p-2t+1}$ | ⋯ | $\gamma^2_{p-3t+3}$ | ⋯ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 0 | -1 | -1 | -1 | 0 | 1 | 1 | 1 | 0 | -1 | 0 |

Table 14: Difference vector $\gamma^3[p, t-2]$.

Proof. We have $\gamma^3_j = \sum_{i=0}^{t-3} \gamma^2_j[p-i,t-1]$. By using the values of $\gamma^2_j[p-i,t-1]$ given in Claim 27, we get that: for $j > p$, $\gamma^2_j = 0$; $\gamma^3_p = 1$; for $p-1 \geq j \geq p-t+3$, $\gamma^3_j = -1+1 = 0$; $\gamma^3_{p-t+2} = 0-1 = -1$; $\gamma^3_{p-t+1} = -1+0 = -1$; $\gamma^3_{p-t} = 0-1 = -1$; for $p-t-1 \geq j \geq p-2t+4$, $\gamma^3_j = 1+0-1 = 0$; $\gamma^3_{p-2t+3} = 1+0 = 1$; $\gamma^3_{p-2t+2} = 0+1 = 1$; $\gamma^3_{p-2t+1} = 1+0 = 1$; for $p-2t \geq j \geq p-3t+4$, $\gamma^3_j = -1+1 = 0$; $\gamma^3_{p-3t+3} = 0-1 = -1$; and for $j < p-3t+3$, $\gamma^3_j = 0$.

Here again a careful but tedious analysis of all the subsequence of deviations indicates that all their coordinates are at least $-3$. However, by using Lemma 20 we can easily prove that this sequence is 4-balanced. In fact, we will prove by induction that $\sum_{i=0}^{r} \overrightarrow{\gamma}^2[p-i,t-1]$ is 4-balanced for any $r \leq t-3$. That is true for $r=0$, as $\overrightarrow{\gamma}^2[p,t-1]$ is 3-balanced. Suppose that $\sum_{i=0}^{r} \overrightarrow{\gamma}^2[p-i,t-1]$ is 4-balanced for some $r \leq t-3$. We apply Lemma 20 with $\overrightarrow{\Phi}^1 = \sum_{i=0}^{r} \overrightarrow{\gamma}^2[p-i,t-1]$ and $\overrightarrow{\Phi}^2 = \overrightarrow{\gamma}^2[p-r-1,t-1]$. By our induction hypothesis, we have $h_1 = 4$. Furthermore, all the coefficients of $\overrightarrow{\Phi}^1$ are greater than $-1$ by Claim 27 when $r=0$ or Claim 28 when $r>0$. Therefore, $\min_i \Phi^1_i = -1$. Finally $\overrightarrow{\Phi}^2$ is 3-balanced and $\sum_{i=0}^{r+1} \overrightarrow{\gamma}^2[p-i,t-1]$ is also $\max(4,3+1) = 4$-balanced. By induction $\overrightarrow{\gamma}^3[p,t-2]$ is 4-balanced. ⋄

Again we can repeat $t$ times the cascades $\overrightarrow{\gamma}^3[p,t-2]$ for creating a sequence of cascades that we call $\overrightarrow{\gamma}^4[p,t] = \sum_{i=0}^{t-1} \overrightarrow{\gamma}^3[p-i,t-2]$. This is enough to prove Theorem 16. Nevertheless, in order to reduce the number of non-zero coordinates to four, we will complete our construction with some trick already used in the previous subsection (Case $k=4$); namely, we will complete $\overrightarrow{\gamma}^3[p,t-2]$ with a sequence of 1-deviations. Note that in doing so, we improve the ratio between the number of 3-deviations and the number of vertices. Thus, after having done $\overrightarrow{\gamma}^3[p,t-2]$, we apply the cascade $\overrightarrow{\alpha}^1[p-t+2,p-3t+3,t-1]$ (see Claim 22). Let $\overrightarrow{\tau}^3[p,t-2]$ be the sequence we obtain. As shown in Claim 29, it only has 4 non-zero coordinates.

**Claim 29.** *For* $3 \le t \le (p+2)/3$ *the coordinates of* $\overrightarrow{\tau}^3[p,t-2] = \overrightarrow{\gamma}^3[p,t-2] + \overrightarrow{\alpha}^1[p-t+2,p-3t+3,t-1]$ *satisfy:* $\tau^3_p = 1, \tau^3_{p-t+1} = \tau^3_{p-t} = -1, \tau^3_{p-2t+1} = 1$, *and* $\tau^3_j = 0$ *for all the other values of* $j$*. Furthermore it is* 4*-balanced.*

Proof. Compared to $\overrightarrow{\gamma}^3[p,t-2]$ only 4 coordinates have been changed. We get $\tau^3_{p-t+2} = \gamma^3_{p-t+2} + \alpha^1_{p-t+2} = -1+1 = 0$, $\tau^3_{p-2t+3} = \tau^3_{p-2t+2} = 1-1 = 0$ and $\tau^3_{p-3t+3} = -1+1 = 0$. All the other coordinates remain the same. To prove that $\overrightarrow{\tau}^3[p,t-2]$ is 4-balanced, we apply Lemma 20 with $\overrightarrow{\Phi}^1 = \overrightarrow{\gamma}^3[p,t-2]$ and $\overrightarrow{\Phi}^2 = \overrightarrow{\alpha}^1[p-t+2,p-3t+3,t-1]$. We have proved that $\overrightarrow{\Phi}^1$ is 4-balanced and by Claim 28 $\min_i \Phi^1_i = -1$. Furthermore, $\overrightarrow{\Phi}^2$ is 1-balanced. Hence, $\overrightarrow{\tau}^3[p,t-2]$ is $\max(4,1+1) = 4$-balanced. ⋄

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

**Claim 30.** *For* $3 \le t \le (p+3)/4$ *the coordinates of* $\overrightarrow{\tau}^4[p,t] = \sum_{i=0}^{t-1} \overrightarrow{\tau}^3[p-i,t-2]$ *satisfy: for* $p \ge j \ge p-t+2$, $\tau^4_j = 1$*; for* $p-t \ge j \ge p-2t+2$, $\tau^4_j = -2$*; for* $p-2t \ge j \ge p-3t+2$, $\tau^4_j = 1$*; and* $\tau^4_j = 0$ *for all the other values of* $j$*. Furthermore it is* 6*-balanced.*

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

We will prove by induction that $\sum_{i=0}^{r} \overrightarrow{\tau}^3[p-i,t-2]$ is 6-balanced for any $r \le t-1$. That is true for $r = 0$ as $\overrightarrow{\tau}^3[p,t-2]$ is 4-balanced. Suppose that $\sum_{i=0}^{r} \overrightarrow{\tau}^3[p-i,t-2]$ is 6-balanced for some $r \le t-2$. We apply Lemma 20 with $\overrightarrow{\Phi}^1 = \sum_{i=0}^{r} \overrightarrow{\tau}^3[p-i,t-2]$ and $\overrightarrow{\Phi}^2 = \overrightarrow{\tau}^3[p-r-1,t-2]$. By our induction hypothesis, we have $h_1 = 6$. Furthermore, all the coefficients of $\overrightarrow{\Phi}^1$ are greater than $-2$, that follows from Claim 29 when $r = 0$ and from Claim 30 when $r > 0$. Hence, $\min_i \Phi^1_i = -2$. Finally, $\overrightarrow{\Phi}^2$ is 4-balanced. Hence, $\sum_{i=0}^{r+1} \overrightarrow{\tau}^3[p-i,t-2]$ is also $\max(6,4+2) = 6$-balanced and by induction $\overrightarrow{\tau}^4[p,t]$ is 6-balanced. ⋄

*Proof.* (**End of the proof of Theorem 16**) The cascade $\overrightarrow{\tau}^4[p,t]$ consists of $t$ cascades $\overrightarrow{\tau}^3[p-i,t-2]$, each of them consisting of a cascade $\overrightarrow{\gamma}^3[p-i,t-2]$ and $(t-1)$ 1-deviations. $\overrightarrow{\gamma}^3[p-i,t-2]$ itself consists of $t-2$ cascades $\overrightarrow{\gamma}^2[p-i,t-1]$, each of them consisting of $t-1$ cascades $\overrightarrow{\gamma}^1[p-i,t-1]$ each of them consisting of $t$ 3-deviations $\overrightarrow{\gamma}[p-i]$. Thus, the cascade $\overrightarrow{\tau}^4[p,t]$ contains $t^2(t-1)(t-2) = \theta(t^4)$ 3-deviations (plus $t(t-1)$ 1-deviations, but that is negligible).

By Claim 30, $\overrightarrow{\tau}^4[p,t]$ is 6-balanced. If we choose as starting partition one with 6 groups of each size $i, 1 \le i \le p-1$, then 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 of any size $i$; for that we choose a specific group of size 1 and sucessively move with $(i-1)$ 1-deviations one element of $i-1$ other groups of size 1 to form a group of size $i$. We do it 6 times for each size $i, 1 \le i \le p-1$. Of course that is possible only if $n \ge 6p(p-1)/2$.

Finally, note that in order for all the coordinates in the cascade to have a meaning we must choose $p$ and $t$ such that $p-4t+3 \ge 0$. Let us choose $p = 4t-3$; then the number of vertices is $n = 6\sum_{i=0}^{p-1} i = 6p(p-1)/2$ $= (48t-36)(t-1) = \Theta(t^2)$.

In summary we have built a cascade which contains $\Theta(t^4) = \Theta(n^2)$ 3-deviations. Therefore, $L(3,n) = \Omega(n^2)$. □

# 6 Conclusion

Our analysis shows that, in the case of a polarized society (*i.e.*, with only friends and enemies), even modest cooperations between $k = 4$ users can surprisingly delay the convergence time of the dynamics of formation of groups. This result was obtained by using an all new combinatorial approach for this problem that may be of independent interest for further studies. In particular, we would find it interesting to know whether our techniques in this paper could be applied to establish the asymptotic worst-case time of convergence for this dynamics, for every $k \ge 3$.

From the complexity point of view, the main open question is to determine the complexity of computing a $k$-stable partition for $k \ge 4$. In particular, is this problem PLS-complete? The reduction presented in [Duc16a] – where it is proved that computing a 1-stable partition is already PTIME-hard – may be helpful in order to answer this open question.

Finally, we intend to study several possible generalizations of the present model of group formation. In [ABK+16], Angel et al. introduced the more general MAX $k$-COLORED CLUSTERING problem, where informally speaking, there are now different categories of information considered (represented by edge-colors). A game-theoretic study of this problem remains to be done. Kleinberg and Ligett also proposed in [KL13] to include in the model different levels of friendship and conflict, as well as the possibility for neutral interactions (represented by positive, negative and zero edge-weights, respectively). In a forthcoming paper, we study this generalization. Our main results there relate the existence of a $k$-stable partition, for any fixed $k$, with the palette of social evaluations (= edge-weights) available.

**Acknowledgments:** We thank the referees of this article for their careful reading and helpful remarks.

# References


[ABK+16] E. Angel, E. Bampis, A. Kononov, D. Paparas, E. Pountourakis, and V. Zissimopoulos. Clustering on k-edge-colored graphs. Discrete Applied Mathematics, 2016.

[Bal04] C. Ballester. NP-completeness in hedonic games. Games and Economic Behavior, 49(1):1–30, 2004.

[BM08] J. A. Bondy and U. S. R. Murty. Graph theory. Grad. Texts in Math., 2008.

[Bry73] T. Brylawski. The lattice of integer partitions. Discrete Mathematics, 6(3):201 – 219, 1973.

[CKPS10] I. Chatzigiannakis, C. Koninis, P. N. Panagopoulou, and P. G. Spirakis. Distributed game-theoretic vertex coloring. In OPODIS'10, pages 103–118, 2010.

[CNS18] J. Chen, R. Niedermeier, and P. Skowron. Stable marriage with multi-modal preferences. In Proceedings of the 2018 ACM Conference on Economics and Computation, EC '18, pages 269–286, New York, NY, USA, 2018. ACM.

[DBHS06] D. Dimitrov, P. Borm, R. Hendrickx, and S-C. Sung. Simple priorities and core stability in hedonic games. Social Choice and Welfare, 26(2):421–433, 2006.

[Duc16a] G. Ducoffe. The parallel complexity of coloring games. In International Symposium on Algorithmic Game Theory, pages 27–39. Springer, 2016.

[Duc16b] G. Ducoffe. Propriétés métriques des grands graphes. PhD thesis, Université Côte d'Azur, December 2016.

[EGM12] B. Escoffier, L. Gourvès, and J. Monnot. Strategic coloring of a graph. Internet Mathematics, 8(4):424–455, 2012.

[FMZ17] M. Flammini, G. Monaco, and Q. Zhang. Strategyproof mechanisms for additively separable hedonic games and fractional hedonic games. In WAOA, pages 301–316, 2017.

[GK86] 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.

[Haj06] J. Hajduková. Coalition formation games: A survey. International Game Theory Review, 8(04):613–641, 2006.

[HJ17] M. Hoefer and W. Jiamjitrak. On proportional allocation in hedonic games. In SAGT, pages 307–319. Springer, 2017.

[HW79] G. H. Hardy and E. M. Wright. An introduction to the theory of numbers. Oxford University Press, 1979.

[JPY88] 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.

[KL13] J. Kleinberg and K. Ligett. Information-sharing in social networks. Games and Economic Behavior, 82:702–716, 2013.

[MS17] M. Mnich and I. Schlotter. Stable marriage with covering constraints–a complete computational trichotomy. In SAGT, pages 320–332. Springer, 2017.

[OBI+17] K. Ohta, N. Barrot, A. Ismaili, Y. Sakurai, and M. Yokoo. Core stability in hedonic games among friends and enemies: impact of neutrals. In Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17, pages 359–365, 2017.

[OM16] I. Olkin and A. W. Marshall. Inequalities: theory of majorization and its applications, volume 143. Academic press, 2016.

[PS08] P. N. Panagopoulou and P. G. Spirakis. A game theoretic approach for efficient graph coloring. In ISAAC'08, pages 183–195, 2008.