Preprints.org

Article

Mathematical Programming for the Dynamics of Opinion Diffusion

Andrea Ellero [*,‡] , Giovanni Fasano [‡] , Daniela Favaretto [‡]

Article

Mathematical Programming for the Dynamics of Opinion Diffusion

Andrea Ellero [1,‡,*] , **Giovanni Fasano [2,‡]** **and Daniela Favaretto [3,‡]**

[1]   Department of Management, Ca' Foscari University of Venice; ellero@unive.it
[2]   Department of Management, Ca' Foscari University of Venice; fasano@unive.it
[3]   Department of Management, Ca' Foscari University of Venice; favaret@unive.it
[*]   Correspondence: ellero@unive.it
[‡]   All the authors contributed equally to this work.

Abstract: The focus of this paper is on analyzing the role and the choice of parameters in sociophysics diffusion models, by leveraging the potentialities of sociophysics from a mathematical programming perspective. We first present a generalised version of Galam's opinion diffusion model (see, e.g. [8,10]). For a given selection of the coefficients in our model, this proposal yields the original Galam's model. The generalised model suggests guidelines for possible alternative selection of its parameters that allow to foster diffusion. Examples of the parameters selection process as steered by numerical optimisation, taking into account various objectives, are provided.

Keywords: Sociophysics; Stochastic Models; Galam's Diffusion Model; Mathematical Programming; Parameter Optimisation

1. Introduction

Opinion diffusion is a fundamental aspect of human interaction, playing a pivotal role in the spread of innovations, and the emergence of consensus. Understanding how opinions spread and evolve within social networks is crucial for comprehending the emergence of collective behavior, the formation of public opinion, and the polarization of societies.

Social sciences provide plenty of applications, where the importance of diffusion dynamics is studied and witnessed, like in marketing (see e.g. [2,12,13]), agent-based modeling (see e.g. [14]) and sociophysics, where the papers of Serge Galam serve as a guide for researchers (see e.g. [9]).

Galam's opinion diffusion model (see [8]) provides a framework for investigating and understanding the dynamics of opinion formation in social systems. Galam's model sheds light into the mechanisms that drive opinion dynamics. In particular, Galam's model describes the formation of collective opinion convergence processes using concepts from statistical physics. To analyze them, along with several generalisations of the model, the authors of the current paper have proposed the alternatives in [4] (showing some drawbacks when applying Galam's model with a relatively small number of agents), and in [5] (where a novel subset of agents, namely the *opinion leaders*, are introduced, to duly represent a possible communication bias). Furthermore, additional literature covering a number of keynote issues can be found in [9–11] and therein references.

The standard Galam's model in [8] assumes that individuals in a group influence and are influenced by their peers. After several iterations within repeated group discussions, each agent may consequently influence the opinion diffusion. Galam's stylized model simplifies the complexity of opinions by considering a binary state, where individuals can adopt either of two possible opinions, thus allowing a straightforward mathematical analysis of the diffusion process.

We suggest a generalisation of Galam's model and then offer evidence in an optimisation framework that some of the diffusion parameters can be successfully adjusted in order to steer the diffusion process. Combining Galam's model with specific Linear Programming (LP) formulations, we take into account both theoretical and numerical results concerning the role of the model's parameters,

in order to better control the dynamics of the diffusion: namely we aim at maximizing the probability that an opinion can spread among agents.

The paper is organised as follows: the essential characteristics of Galam's model [8] are recapped in Section 2. Then, in Section 3, we present and discuss a generalisation of Galam's model. The optimisation of the parameters of the new model via linear programming allows to speed up opinion diffusion at a fixed step of the process: the optimisation problem is discussed in Section 4 while Section 5 provides numerical examples of its resolution. Section 6 discusses a dynamic programming extension of the optimisation problem that optimizes the values of the parameters along the whole diffusion process, till convergence. Section 7 contains final remarks and concludes the paper.

2. Recap of Galam's model

This section reviews the foundational elements of Galam's model in [8]. Think of a group of N people (*agents*), each of whom may hold one of two opposing views (let's say, '+' or '−'), regarding a specific issue. These agents meet, during several repeated rounds, within subgroups of smaller cardinality, to converse and possibly change one another's minds.

Each agent at round (*time*) t has a probability a_k of being a member of a subgroup of size k, being L the maximum allowed size of groups. In the original formulation of Galam's model the values $a_1, \ldots, a_L$ are exogenous parameters such that

$$\sum_{k=1}^{L} a_k = 1, \qquad a_k \geq 0, \quad k = 1, \ldots, L.$$

Joining a group discussion at time t, any agent may, at the beginning of the subsequent time $t + 1$, modify their position (e.g. '+' becomes '−' or viceversa) in accordance with a unique law: the majority rule. Indeed, all agents in a subgroup take the position of the majority in that group in the end of time t. The rule for reversing opinion is slightly biased in favor of the negative opinion '−', since tie breaks in favor of '−'. This may lead to a substantial bias in case the subgroup has an even number of members.

Calling $P_+(t)$ the probability that an agent opinion is '+' at time t, the probability of opinion '−' is then $P_-(t) = 1 - P_+(t)$. The probability $P_+(t + 1)$ is estimated in [8] as

$$P_+(t + 1) = \sum_{k=1}^{L} a_k \sum_{j=\left\lfloor \frac{k}{2}+1 \right\rfloor}^{k} C_j^k P_+(t)^j \{1 - P_+(t)\}^{k-j}, \tag{1}$$

where $\lfloor z \rfloor$ is the largest integer less or equal to z, and C_j^k is the binomial coefficient $\binom{k}{j}$.

Let $N_+(t)$ ($N_-(t) = N - N_+(t)$) be the number of agents who at time step t have opinion '+' ('−'). Observe that for small values of N, setting $P_+(0) = N_+(0)/N$, where $N_+(0)$ is the number of agents thinking '+' at time $t = 0$, for any $t \geq 1$ the quantity $P_+(t)$ may possibly differ from the 'actual' frequency of '+', i.e. $N_+(t)/N$ (see [4]). Nevertheless, the estimated probability $P_+(t + 1)$ always lies in the interval $[0, 1]$ being

$$\sum_{j=\left\lfloor \frac{k}{2}+1 \right\rfloor}^{k} C_j^k P_+(t)^j \{1 - P_+(t)\}^{k-j} < \sum_{j=0}^{k} C_j^k P_+(t)^j \{1 - P_+(t)\}^{k-j} = 1.$$

A relevant role in the model is taken by the value $\hat{P}_+$ such that

$$\text{when } P_+(0) > \hat{P}_+ \text{ then } \lim_{t \to \infty} P_+(t) = 1,$$
$$\text{when } P_+(0) < \hat{P}_+ \text{ then } \lim_{t \to \infty} P_+(t) = 0,$$
$$\text{when } P_+(0) = \hat{P}_+ \text{ then } P_+(t) = P_+(0), \qquad \forall \, t > 0.$$

$\hat{P}_+$ constitutes the threshold value such that when $N_+(0)/N$ is greater than $\hat{P}_+$, then all agents will eventually have opinion '+'. Conversely, when $t \to \infty$ all the agents will definitely think '−' if

$N_+(0)/N < \hat{P}_+$. The value $\hat{P}_+$ is addressed as the *killing point* of the model. In Figure 1 we can observe the dynamics described by relation (1), when the largest cardinality L of the subgroups takes integer values in the interval $[3,20]$ and $a_k = 1/L$, for any k: for each of the 18 curves, the killing point is identified by the intersection with the bisector line other than the extreme points $(0,0)$ and $(1,1)$. Observe that for small values of L no killing point appears (or equivalently it coincides with the extreme point $(1,1)$). Figure 2 reports a zoom perspective of Figure 1.

Figure 1. Dynamic curves described by relation (1), when the largest cardinality L of the subgroups takes integer values in the interval $[3,20]$ and $a_k = 1/L$, for any k.
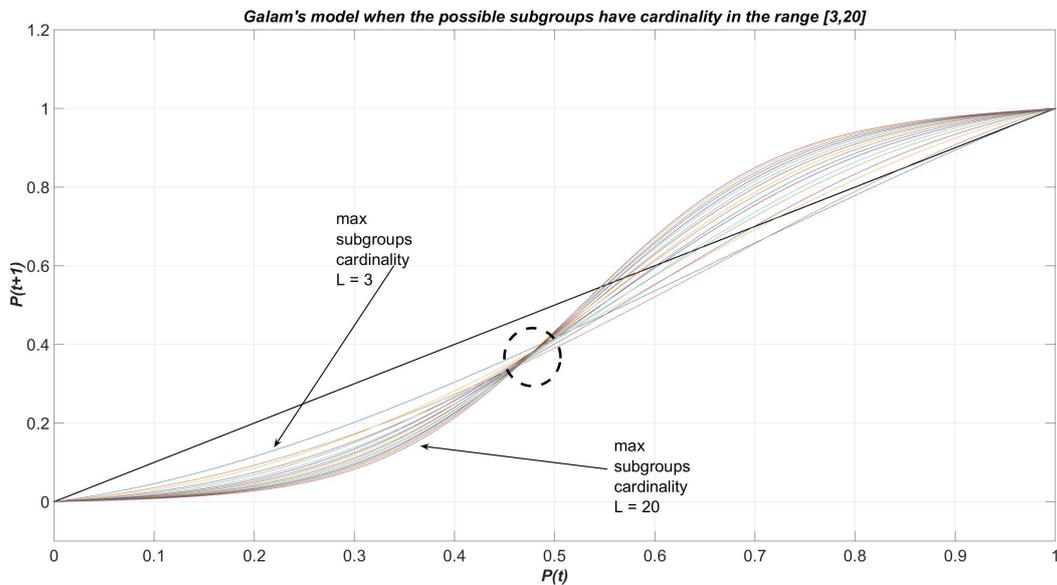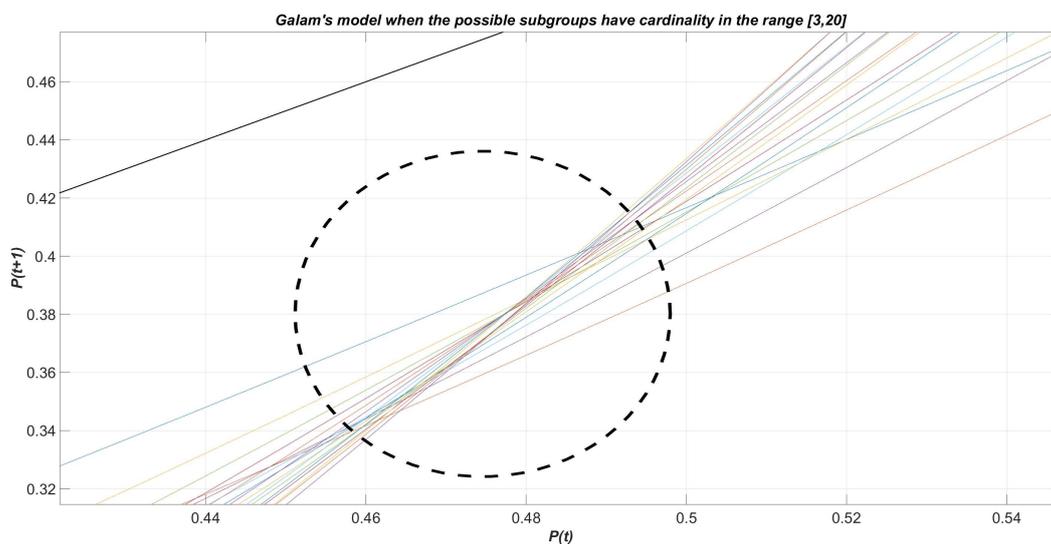

Figure 2. Dynamic curves described by relation (1), when the largest cardinality L of the subgroups takes integer values in the interval $[3,20]$ and $a_k = 1/L$, for any k: zoom of the area of intersections among curves.

3. Capturing additional dynamics using a generalised Galam's model

According to model (1), a strict majority of members with positive opinion is a necessary and sufficient condition to have opinion $+$ for all the subgroup members.

Since the ultimate choice in a group is frequently the product of a discussion, rather than a simple tally of the two opposing viewpoints in the group, a rule that is so rigorous yet so straightforward could not correctly reflect the actual opinion dynamics in practice. As an illustration, think about the propagation of political opinions. The probability that everyone in the group would share the same view (e.g. $+$) at the close of a conversation may rise *smoothly* with the number of members who had that opinion prior to the meeting. As a second application, the reader may consider the audience of social networks, where the dynamics of information spreading was widely investigated in a dedicated literature (see, e.g. [1]).

Accordingly, we suggest to generalise Galam's model (1) taking into account a softer dynamics of opinion change, i.e., without using rigid majority as a necessary and sufficient condition to shift the entire group to the same view. In particular, we assume that after the conversations, all agents in a group will have opinion $+$ with a probability α_j^k that is dependent on the size k of the group and the number j of agents that had opinion $+$ before the discussion. We use the following definition to model the relationship between the probabilities of having an opinion $+$ before (i.e. $P_+(t)$) and after (i.e. $P_+(t+1)$) discussion (see [3] for a preliminary version of the model):

$$P_+(t+1) = \sum_{k=1}^{L} a_k \sum_{j=0}^{k} \alpha_j^k C_j^k P_+(t)^j \{1 - P_+(t)\}^{k-j}. \tag{2}$$

Model (2) generalises (1), since it is possible to select indeed the probabilities $\{\alpha_j^k\}$ that exactly replicate model (1) (see also Figure 3):

$$\alpha_j^k = \begin{cases} 0 & \text{if } j < \left\lfloor \dfrac{k}{2} + 1 \right\rfloor \\[2ex] 1 & \text{if } j \geq \left\lfloor \dfrac{k}{2} + 1 \right\rfloor \end{cases} \tag{3}$$

Figure 3. For a given $1 \leq k \leq L$ the model (3) considers the above *step-shaped* choice for the coefficients $\{\alpha_j^k\}$. In particular, if $j < \lfloor k/2 + 1 \rfloor$ then $\alpha_j^k = 0$, otherwise $\alpha_j^k = 1$, so that this choice corresponds exactly to obtain the original Galam's model (1).

As mentioned, probabilities $\{\alpha_j^k\}$ should reasonably increase with the number j of $+$ in a subgroup, and decrease with respect to the size of the group k. In this regard, it is reasonable to expect for the

probability α_j^k to become negligible when the number j of $+$ in a subgroup is relatively small (i.e., much lower than the strict majority), and to become close to 1 as the number of $+$ noticeably increases (i.e., higher than strict majority). In general, the probability that all agents in the subgroup will adopt the opinion $+$ is conceivably a function rising from zero to one as j increases. If we assume that this function of j is linear, we may precisely define the probabilities α_j^k as (see Figure 4)

$$
\alpha_j^k = \begin{cases}
0 \,, & \text{if } j < \left\lfloor \frac{k}{2} + 1 \right\rfloor + z_j^k - h \\[2mm]
\dfrac{1}{2} + \dfrac{\left[j - \left(\left\lfloor \frac{k}{2} + 1 \right\rfloor + z_j^k \right) \right]}{2h} \,, & \text{if } \left\lfloor \frac{k}{2} + 1 \right\rfloor + z_j^k - h \leq j < \left\lfloor \frac{k}{2} + 1 \right\rfloor + z_j^k + h \\[2mm]
1 \,, & \text{if } j \geq \left\lfloor \frac{k}{2} + 1 \right\rfloor + z_j^k + h
\end{cases}
\tag{4}
$$

When $z_j^k = h = 0$, the probabilities defined in (4) are exactly the same as in (3) . The value of z_j^k in (4) represents a shift (either positive or negative) from the value $\lfloor \frac{k}{2} + 1 \rfloor$, while the slope of the ramp in Figure 4 is $1/(2h)$.

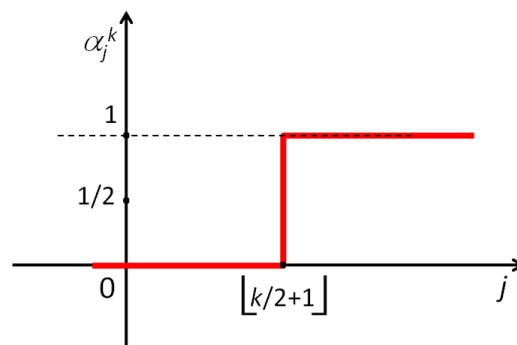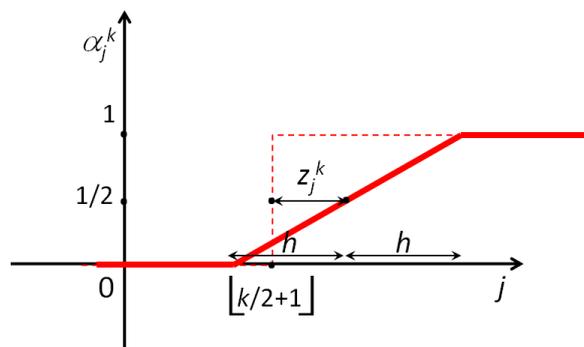**Figure 4.** For a given $1 \leq k \leq L$ the model (4) considers the above *piecewise-linear* choice for the coefficients $\{\alpha_j^k\}$, where z_j^k represents a shift with respect to the abscissa $\lfloor k/2 + 1 \rfloor$, and $1/(2h)$ is the slope of the ramp.

To better grasp the geometric insight behind the choice of the parameters in (4), let us consider the simple numerical example reported in [8], where $L = 4$, $a_1 = 0$, $a_2 = a_3 = a_4 = 1/3$ (and $h = z_j^k = 0$ for all j and k). As reported in [8] the killing point (KP) lies in between 0.84 and 0.87 (see Figure 5). Then, we also have in Figure 6 the model (2) with the choice (4), after setting $h = z_j^k = 0.5$ for any j and k. We can immediately infer that the dynamics of Galam's model in [8] is definitely spoiled when the choice for the coefficients $\{\alpha_j^k\}$ in (3) is replaced by (4). Equivalently, even in case $P_+(t) = 1$ (complete consensus among agents to think $+$) then the choice (4) imposes a *regression* towards the stable point given by the origin.
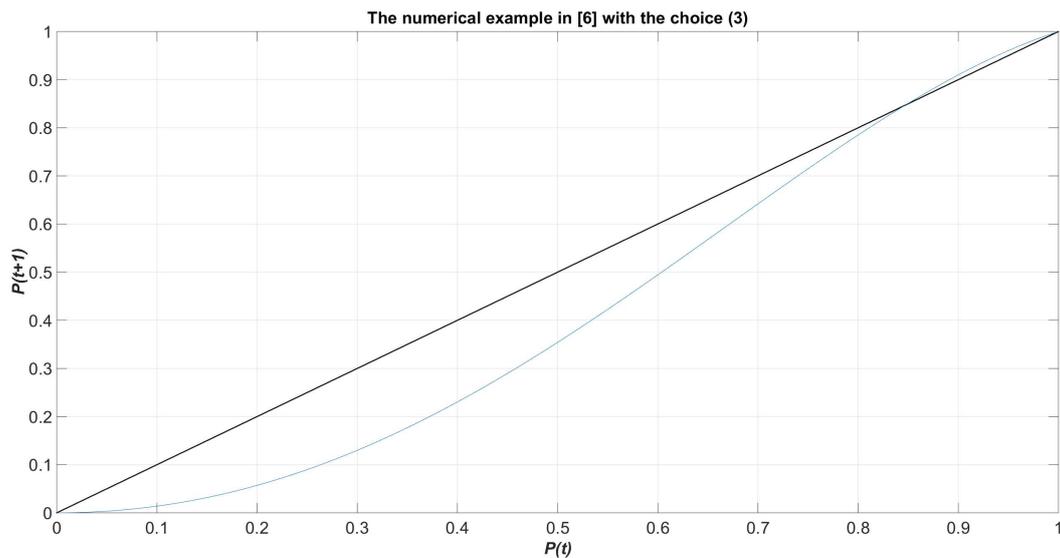
Figure 5. The simple numerical example reported in [8], where $L = 4$, $a_1 = 0$, $a_2 = a_3 = a_4 = 1/3$.
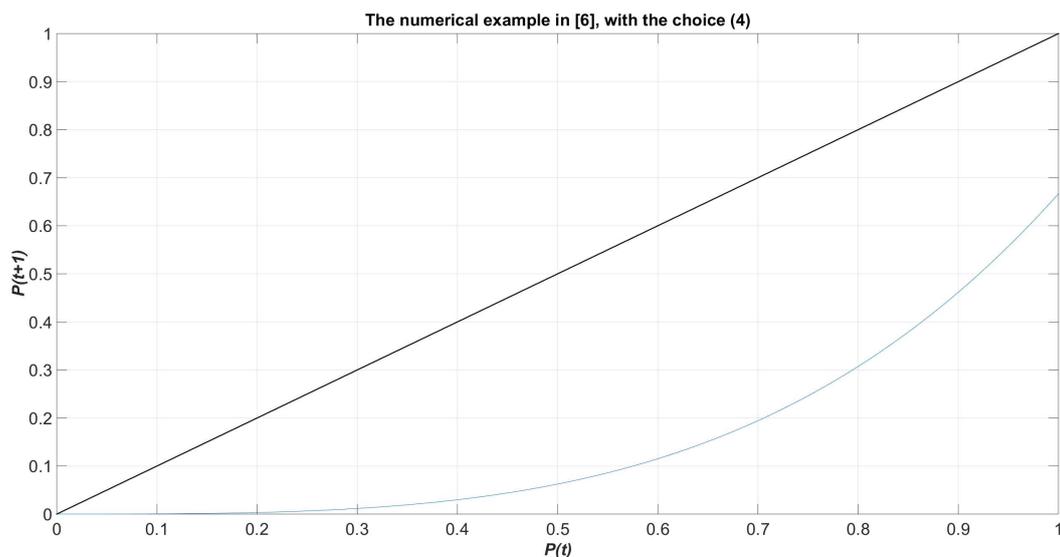
Figure 6. The simple numerical example reported in [8], where $L = 4$, $a_1 = 0$, $a_2 = a_3 = a_4 = 1/3$, and now the choice (3) is replaced by the selection (4).

Hence, the presence of a killing point for Galam's model seems to be strictly related to the expression (1), and introducing the coefficients $\{\alpha_j^k\}$ definitely spoils also the killing point. Moreover, we observe that, in case the majority rule is not explicitly applied in (1), then the dynamics of the model is strongly altered.

Nevertheless, the model (2) retains a number of properties and can suggest fruitful results, following both the next guidelines:

- it provides clues on the perspective for *studying* the process of diffusion of information;
- it may suggest proper ways to *control* information spreading among agents.

At least a couple of intriguing issues in applications can steer the choice of the coefficients $\{\alpha_j^k\}$, ranging from small–medium to large scale (number of agents) applications:

- **(One Period Analysis)** we may assess values for $\{\alpha_j^k\}$ so that, considering at time t the probability $P_+(t)$, we maximize the probability $P_+(t+1)$ at time $t+1$. In this regard, the values $\{\alpha_j^k\}$ represent *costs* for steering the information among agents and fostering the positive opinion: the larger $\{\alpha_j^k\}$ (i.e. higher costs) the larger the probability to convey a group to opinion '+'. Of course the combination of the proper values for each of the coefficients $\{\alpha_j^k\}$ is strongly related to the value $P_+(t)$, the probabilities $\{a_k\}$ and the binomial coefficients C_j^k;

- **(Multi-Period Analysis)** similarly to the previous item we may decide to select $\{\alpha_j^k\}$ such that, considering again at time t the probability $P_+(t)$, we maximize the probability $P_+(T)$ at time $T > t$, and compute the entire sequence $P_+(t+1), P_+(t+2), \dots, P_+(T)$ of intermediate probabilities.

The last two scenarios can be declined in a variety of contexts, including production and Marketing, where consumers play the role of agents and the aforementioned perspectives may be the answer to the question: what is (if any) the necessary extra effort in terms of advertising campaign, in order to promote a certain spread of the products or customers' expectations on items ? In this regard the quantities $\{\alpha_j^k\}$ summarise the resulting effort, while from a mathematical programming perspective they represent unknowns to be assessed. This also suggest that a number of additional questions may arise adopting the model (2), mapping relevant practical instances where a tentative forecast is sought in order to plan the use of (possibly scarce) resources to control the process of diffusion of the information. For the sake of brevity, we itemize below some proposals that are representative but not exhaustive:

- **(Feasibility Problem)** given the value of $P_+(t)$ and the target value $\bar{P}$, we may determine a set of values for the coefficients $\{\alpha_j^k\}$ such that, starting at t from $P_+(t)$, at time $T > t$ we will have $P_+(T) \geq \bar{P}$;

- **(Focus Group Problem)** given the value of $P_+(t)$ and the target value $\bar{P}$ and the time $T > t$, we may be interested to determine a set of values for the coefficients $\{\alpha_j^k\}$ such that $\alpha_m^{(k)} \leq \alpha_j^k \leq \alpha_M^{(k)}$, where α_m, α_M are possible bounds, and $P_+(T) \geq \bar{P}$. This scenario models a specific interest on those subgroups of cardinality k.

4. Single period maximization of diffusion

As from the analysis in Section 3, one of the possible proposals to couple the dynamics described through the model (2) with a single period optimization framework, is given by the following LP scheme:

$$\max_{\alpha} \quad \sum_{k=1}^{L} a_k \sum_{j=0}^{k} \alpha_j^k C_j^k P_+(t)^j \{1 - P_+(t)\}^{k-j}, \tag{5}$$

$$\alpha_j^k \leq \alpha_{j+1}^k \qquad j = 0, 1, \dots, k-1; \ k = 1, \dots, L, \tag{6}$$

$$\alpha_j^k \geq \alpha_j^{k+1} \qquad j = 0, 1, \dots, k; \ k = 1, \dots, L-1, \tag{7}$$

$$\sum_{k=1}^{L} \sum_{j=0}^{k} \alpha_j^k \leq b(t) \tag{8}$$

$$0 \leq \alpha_j^k \leq 1 \qquad j = 0, 1, \dots, k; \ k = 1, \dots, L. \tag{9}$$

The last optimization problem assumes that the value $P_+(t)$ is given, so that both the objective function and the constraints are linear in the set $\{\alpha_j^k\}$. Though the meaning associated with the objective function is relatively clear, for the constraints some clarifications seem necessary, indeed:

- the constraints (6) state that for a given dimension k of a subset, larger values of the number of agents thinking $+$ imply a larger probability α_j^k;
- similarly in the constraints (7), when the number of agents thinking $+$ remains constant, the larger the set cardinality (i.e. k) the smaller the probability α_j^k;
- the constraint (8) expresses the *overall cost* (i.e. it is a budget constraint) allowed to assess the probabilities $\{\alpha_j^k\}$, at time t;
- since the quantities $\{\alpha_j^k\}$ need to represent *probabilities*, they are required to fulfill also the constraint (9).

In order to avoid either infeasible or straightforward solutions for the optimization problem (5)–(9), the next lemmas gives some indications on the choice of the parameter $b(t)$, for any t (for a proof see [3]).

Lemma 1. *Let be given the linear program* (5)–(9) *with the choice* ($1 \leq k \leq L$ *and* $0 \leq j \leq k$)

$$
\alpha_j^k = \begin{cases} 0, & \text{for any } j < \left\lfloor \frac{k}{2} + 1 \right\rfloor, \\[2ex] 1, & \text{for any } j \geq \left\lfloor \frac{k}{2} + 1 \right\rfloor. \end{cases} \tag{10}
$$

Then, relations (10) *satisfy the constraints* (6)–(7) *and* (9). *Moreover, if* (10) *also satisfies* (8), *then the value of the objective function in* (5) *coincides with* $P_+(t+1)$ *in* (1).

Note that Lemma 1 gives a clear idea of the fact that the solution of the linear program (5)–(9) is substantially nothing else but a generalization of (1). Furthermore, the next result will give a precise indication on the possible values for the budget $b(t)$ in (8).

Lemma 2. *Let be given the optimization problem* (5)–(9) *and let* $\{\alpha_j^k\}$ *be assigned as in* (10), *for any* $1 \leq k \leq L$ *and* $0 \leq j \leq k$. *Then*

$$
\sum_{k=1}^{L} \sum_{j=0}^{k} \alpha_j^k = \begin{cases} \left(\dfrac{L}{2}\right)^2 + \dfrac{L}{2}, & \text{when } L \text{ is even}, \\[3ex] \left(\dfrac{L-1}{2}\right)^2 + L, & \text{when } L \text{ is odd}. \end{cases} \tag{11}
$$

Note that the right sides of (11) explicitly give some hints for the selection of the parameter $b(t)$ in (8), both either in case L is even or it is odd.

As regards some technicalities associated with the solution of the optimization problem (5)–(9), we observe that it is a continuously differentiable problem, where all the functions are linear. Hence, it is a convex problem, so that the set of all its solutions (possibly an empty set or a singleton) is a convex set. This implies that for any given pair of its solutions, all the points in the segment joining them will be solutions, too. Moreover, being a linear program, all its (possible) solutions will lie on vertices of the feasible set, and not in the interior of the feasible set. As well known, this makes its solution relatively simple (throughout any solver based on the Simplex method) and achievable in a polynomial time. Hence, large scale instances are easily tractable and allow for a scalable solution in a number of applications with a great number of agents (e.g. problems involving social media or large social groups).

We complete this section highlighting that one further generalization of the linear program (5)–(9) suggests to disaggregate the budget constraint and to replace the inequality (8) by the set of constraints

$$
\sum_{j=0}^{k} \alpha_j^k \leq b_k(t), \qquad k = 1, \ldots, L.
$$

Observe that the value $b_k(t)$, for any k, represents a budget devoted to possibly affect the solution after working uniquely on the subsets of cardinality k. This increases the flexibility of the model without compromising its overall complexity (being the resulting model yet a linear program).

5. Preliminary numerical experience

This section is devoted to provide a numerical experience where the potentialities of the formulation (5)–(9) are exploited. In this regard, we limit our investigation to a couple of meaningful instances, being the first one a small scale (number of unknowns) problem, while the second one attempts to scale the first one scale by 100.

5.1. Small scale instance

As regards the value of the coefficients in (5)–(9), we consider for the small scale instance the following setting:

$$
\begin{cases}
L = 10 \\
\{a_k\} = \{0.1, 0, 0.15, 0, 0.3, 0, 0.3, 0, 0.15, 0\} \\
P_+(t) = 0.65.
\end{cases}
$$

The above choice corresponds to possibly allow only subgroups of dimension 1, 3, 5, 7 and 9, with probabilities that include null values and possibly repeated values. Furthermore, we set the initial percentage of the population thinking '+' to $P_+(t) = 65\%$. Finally, in order to avoid an empty feasible set in (5)–(9) we also selected respectively the values $b(t) = 5$ and $b(t) = 10$, based on Lemma 2. The outcomes of our numerical experience are obtained using the solver CPLEX 20.1.0.0 available on the platform `https://neos-server.org/neos`. We remark that CPLEX is definitely among the best and faster solvers for LP, along with the alternative solver BARON. Our numerical experience on the above two instances is summarized as follows:

- $b(t) = 5$: the problem (5)–(9) was coded using AMPL (A Mathematical Programming Language, see [6] and [7]). The *presolve* tool in CPLEX eliminated 65 constraints from the formulation, so that the simplified problem reduced to 65 variables and 110 linear inequality constraints. As regards the final solution, CPLEX performed 13 *dual simplex iterations* (i.e. iterations of a method for LP which represents the counterpart of the Simplex method, after exploiting the duality theory), with an overall time for the solution not larger than a couple of seconds. Finally, the value of $P_+(t+1)$ found by the solver was

$$
P_+(t+1) = 0.245428625
$$

with the set of variables $(\alpha_j^k)^*$ given by

	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10
j=0	0	0	0	0	0	0	0	0	0	0
j=1	1	0	0	0	0	0	0	0	0	0
j=2	.	0	0	0	0	0	0	0	0	0
j=3	.	.	2/3	2/3	2/3	0	0	0	0	0
j=4	.	.	.	2/3	2/3	0	0	0	0	0
j=5	.	.	.	.	2/3	0	0	0	0	0
j=6	.	.	.	.	.	0	0	0	0	0
j=7	.	.	.	.	.	.	0	0	0	0
j=8	.	.	.	.	.	.	.	0	0	0
j=9	.	.	.	.	.	.	.	.	0	0
j=10	.	.	.	.	.	.	.	.	.	0

- $b(t) = 10$: again the problem (5)–(9) was coded using AMPL and the *presolve* tool in CPLEX eliminated 65 constraints from the formulation, so that again the simplified problem reduced to 65 variables and 110 linear inequality constraints. CPLEX performed 18 *dual simplex iterations*

with an overall time for computation similar to the case where it was $b(t) = 5$. Finally, the value of $P_+(t+1)$ found by the solver was

$$P_+(t+1) = 0.4392275888$$

with the set of variables $(\alpha_j^k)^*$ given by the matrix

	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10
j=0	1	0	0	0	0	0	0	0	0	0
j=1	1	0	0	0	0	0	0	0	0	0
j=2	.	0	0	0	0	0	0	0	0	0
j=3	.	.	1	1	1	0	0	0	0	0
j=4	.	.	.	1	1	0.285714	0.285714	0	0	0
j=5	.	.	.	.	1	0.285714	0.285714	0	0	0
j=6	.	.	.	.	.	0.285714	0.285714	0	0	0
j=7	.	.	.	.	.	.	0.285714	0	0	0
j=8	.	.	.	.	.	.	.	0	0	0
j=9	.	.	.	.	.	.	.	.	0	0
j=10	.	.	.	.	.	.	.	.	.	0

Note that allowing a larger value of the budget (say $b(t) = 10$), the solver was able to select a larger number of nonzero unknowns, so that the final value of $P_+(t+1)$ was almost doubled with respect to the case $b(t) = 5$. In real applications this turns to imply that, in order to obtain almost twice the effect of the previous case, on this instance twice the value of the budget must be allowed.

5.2. Large scale instance

Here we focus on the same instance of Section 5.1, where conversely a larger number of unknowns is considered. The novel setting of the parameters for the solver are given as follows:

$$\begin{cases} L = 1000 \\ a_k = 1/L, & k = 1, \ldots, L, \\ P_+(t) = 0.65. \end{cases}$$

The above choice corresponds to possibly allow all the subgroups with the same probability, and again the initial percentage of the population thinking '+' is equal to $P_+(t) = 65\%$. Similarly to the small scale instance, we considered two values for the budget, being $b(t) = 50$ and $b(t) = 100$. The computation is again performed adopting the solver CPLEX 20.1.0.0 on the Neos Server. The outcomes of the resulting two instances are summarized as follows:

- $b(t) = 50$: the problem (5)–(9) was coded using AMPL and the *presolve* tool in CPLEX eliminated 501500 constraints from the formulation. The overall simplified problem contained 501500 variables and 1001000 linear inequality constraints. CPLEX performed 192 *dual simplex iterations* with a time of computation smaller than 20 seconds, showing that the linear model can easily scale the time of computation. The final value of $P_+(t+1)$ found by the solver was

  $$P_+(t+1) = 0.01061571566$$

  with the set of variables $(\alpha_j^k)^*$ given by

	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10	k=11	k=12	k≥13
j=0	1	0	0	0	0	0	0	0	0	0	0	0	0
j=1	1	1	1	0	0	0	0	0	0	0	0	0	0
j=2	.	1	1	1	1	0	0	0	0	0	0	0	0
j=3	.	.	1	1	1	1	1	0	0	0	0	0	0
j=4	.	.	.	1	1	1	1	1	0	0	0	0	0
j=5	.	.	.	.	1	1	1	1	1	1	0	0	0
j=6	.	.	.	.	.	1	1	1	1	1	1	0	0
j=7	.	.	.	.	.	.	1	1	1	1	1	0.8333	0
j=8	.	.	.	.	.	.	.	1	1	1	1	0.8333	0
j=9	.	.	.	.	.	.	.	.	1	1	1	0.8333	0
j=10	.	.	.	.	.	.	.	.	.	1	1	0.8333	0
j=11	.	.	.	.	.	.	.	.	.	.	1	0.8333	0
j=12	.	.	.	.	.	.	.	.	.	.	.	0.8333	0
j=13	.	.	.	.	.	.	.	.	.	.	.	.	0
j=14	.	.	.	.	.	.	.	.	.	.	.	.	0
j>14	.	.	.	.	.	.	.	.	.	.	.	.	0

- $b(t) = 200$: the outcomes have been much similar to the case where $b(t) = 50$, being 501500 the overall number of variables and 1001000 the linear inequality constraints. CPLEX performed 30854 *simplex iterations* with a time of computation smaller than 20 seconds, showing again that the linear model can easily scale the time of computation. The final value of $P_+(t+1)$ found by the solver was now

$$P_+(t+1) = 0.0242061448$$

and we do not report the value of the unknowns for the sake of simplicity.

Observe that in both the last two examples, given that subgroups are allowed to be very large, despite relatively high budget values, the levels reached by the probabilities $P_+(t+1)$ are very low.

As a general achievement, we can easily describe the pattern of the nonzero unknowns which was respected in all the four numerical tests reported above. In particular, we can easily prove that *the solution of the formulation* (5)–(9) *only fills the principal submatrix of unknowns of order $h \times h$*, depending h on the value of the budget parameter along with the number of possible subgroups L. In the end, this also allows to easily assess a number of null unknowns, and possibly reduce the complexity of the overall formulation.

6. A Multi-Period Model

This section considers a multi-period reformulation of the Linear Program (5)-(9), in order to further generalize the model (2). In particular, in place of the *single-period* model (5)-(9), where at time step t the quantity $P_+(t+1)$ is maximized starting from $P_+(t)$, we can consider a *multi-period* approach

with the ultimate goal of maximizing $P_+(T+1)$, being T the time horizon. The latter approach is summarized in the next Nonlinear Program:

$$\max_{\alpha} \quad P_+(T+1), \tag{12}$$

$$P_+(t+1) = \sum_{k=1}^{L} a_k \sum_{j=0}^{k} \alpha_j^k(t) C_j^k P_+(t)^j (1 - P_+(t))^{k-j}, \qquad t = 1, \ldots, T, \tag{13}$$

$$\alpha_j^k(t) \leq \alpha_{j+1}^k(t) \qquad t = 1, \ldots, T; \ j = 0, 1, \ldots, k-1; \ k = 1, \ldots, L, \tag{14}$$

$$\alpha_j^k(t) \geq \alpha_j^{k+1}(t) \qquad t = 1, \ldots, T; \ j = 0, 1, \ldots, k; \ k = 1, \ldots, L-1, \tag{15}$$

$$\sum_{k=1}^{L} \alpha_j^k(t) \leq b_j(t) \qquad t = 1, \ldots, T; \ j = 0, 1, \ldots, k, \tag{16}$$

$$0 \leq \alpha_j^k(t) \leq 1 \qquad t = 1, \ldots, T; \ j = 0, 1, \ldots, k; \ k = 1, \ldots, L. \tag{17}$$

The previous Nonlinear Program includes the unknowns

$$\alpha_j^k(t), \qquad t = 1, \ldots, T; \ j = 0, 1, \ldots, k; \ k = 1, \ldots, L,$$

which represent a larger number of variables with respect to (5)-(9). This implies that the larger the time horizon T, the larger the number of unknowns of the resulting problem formulation. Moreover, as already remarked, (12)-(17) is a Nonlinear Program, which in general increases the difficulty of its solution. In particular, the nonlinearities in (13) might yield a nonconcave overall maximization problem, which implies that a certain number of local maxima (which are not global) might arise. In addition, unlike (5)-(9), the feasible set of (12)-(17) is no more a polyhedron, so that the final solutions not necessarily are located on the boundary of the feasible set. As a consequence, the choice of the starting value $P_+(1)$ in (13) becomes a key factor for at least a couple of reasons:

- the sequence $\{P_+(t)\}$ strongly depends on $P_+(1)$, in a similar fashion of the single-period formulation (5)-(9);
- the choice of $\{P_+(t)\}$ strongly affects the local maximum provided by the solver adopted for (12)-(17).

As regards the constraints (13), observe that they simply state a recursion for the function in (2), at each time step. Moreover, the constraints (14), (15) and (17) have a similar meaning of the constraints (6), (7) and (9), for the single-period formulation. Finally, the budget constraints (16) have a possible double formulation, due to the multi-period scheme.

Proposition 1. *Let the sequence $\{\alpha_j^k(t)\}$ be a global solution of (5)-(9), for any $t = 1, \ldots, T$. Then, $\{\alpha_j^k(t)\}_{t=1,\ldots,T}$ is a global solution of (12)-(17).*

Proof. Observe that from relation (2) we have the following equalities

$$
P_+(t+1) - P_+(t) =
$$
$$
\sum_{t=2}^{T+1} \left[\sum_{k=1}^{L} a_k \sum_{j=0}^{k} \alpha_j^k(t) C_j^k P_+(t)^j (1 - P_+(t))^{k-j} - \sum_{k=1}^{L} a_k \sum_{j=0}^{k} \alpha_j^k(t-1) C_j^k P_+(t-1)^j (1 - P_+(t-1))^{k-j} \right] =
$$
$$
\sum_{t=2}^{T+1} \Delta P_+(t).
$$

Thus, considering that $P_+(1) \in \Re$ is a constant value, the maximization in (12) is equivalent to the maximization

$$
\max \sum_{t=2}^{T+1} \Delta P_+(t). \tag{18}
$$

Moreover, observe that if $P_+(t)$ is given, $\Delta P_+(t+1)$ depends only on the unknowns

$$
\alpha_j^k(t), \qquad j = 0, \ldots, k;\ k = 1, \ldots, L,
$$

and the maximization in (5) is equivalent to

$$
\max_{\alpha} \Delta P_+(t+1).
$$

Now, suppose the sequence $\{\alpha_j^k(t)\}_{j=0,\ldots,k;\ k=1,\ldots,L}$ is a solution of (5)-(9); then, it also satisfies (13) for t, as well as (14), (15), (16) and (17). Thus, if the sequences

$$
\left\{ \alpha_j^k(t) \right\}_{j=0,\ldots,k;\ k=1,\ldots,L} \qquad\qquad t = 1, \ldots, T \tag{19}
$$

solve the problem (5)-(9), for $t = 1, \ldots, T$, then by (18) the values (19) are a feasible point of (12)-(17), satisfying also (18), i.e. (19) is a solution of (12)-(17). $\square$

Lemma 3. *Let* $\left\{ \alpha_j^k(t) \right\}$ *be a global solution of* (5)-(9)*, for any* $t \in \{1, 2, \ldots, T\}$*, where the inequality* (8) *is replaced by the nonlinear inequality*

$$
\phi_t \left[\sum_{k=1}^{L} \alpha_j^k(t) \right] \le b_j(t), \qquad \phi_t : \Re^{\frac{(L+1)(L+2)}{2} - 1} \to \Re.
$$

Then, $\left\{ \alpha_j^k(t) \right\}_{t=1,\ldots,T}$ *is also a global solution of* (12)-(17)*, with* (16) *replaced by*

$$
\phi_t \left[\sum_{k=1}^{L} \alpha_j^k(t) \right] \le b_j(t), \qquad t = 1, \ldots, T;\ j = 0, \ldots, k. \tag{20}
$$

Proof. The proof trivially follows the guidelines of the proof for Proposition 1. $\square$

Lemma 4. *Let* $\left\{ \alpha_j^k(t) \right\}$ *be a global solution of* (5)-(9)*, for any* $t \in \{1, 2, \ldots, T\}$*, where the inequality* (8) *is replaced by the nonlinear inequality*

$$
\Phi \left[\sum_{k=1}^{L} \alpha_j^k(t) \right] \le b_j(t), \qquad \Phi : \Re^{\frac{(L+1)(L+2)}{2} - 1} \to \Re,
$$

and Φ is any convex function. Then, $\left\{\alpha_j^k(t)\right\}_{t=1,\ldots,T}$ is also a feasible solution of (12)-(17)*, with* (16) *replaced by*

$$\Phi\left[\sum_{t=1}^{T}\lambda_t\sum_{k=1}^{L}\alpha_j^k(t)\right] \le b_j(t), \qquad j=0,\ldots,k;\ 0\le\lambda_t\le 1;\ \sum_{t=1}^{T}\lambda_t=1, \tag{21}$$

being $b_j \ge \sum_{t=1}^{T}\lambda_t b_j(t)$.

Proof. The convexity of Φ and the hypotheses imply

$$\Phi\left[\sum_{t=1}^{T}\lambda_t\sum_{k=1}^{L}\alpha_j^k(t)\right] \le \sum_{t=1}^{T}\lambda_t\Phi\left[\sum_{k=1}^{L}\alpha_j^k(t)\right] \le \sum_{t=1}^{T}\lambda_t b_j(t) \le b_j,$$

so that $\left\{\alpha_j^k(t)\right\}_{t=1,\ldots,T}$ also satisfy (21). The rest of the proof follows the guidelines of the proof for Proposition 1. $\square$

Remark 1. *Note that from Lemma* 4*, if the sequences $\left\{\alpha_j^k(t)\right\}$, $t=1,\ldots,T$, are global solutions of* (5)-(9) *for $t=1,\ldots,T$, then the sequence $\left\{\alpha_j^k(t)\right\}_{t=1,\ldots,T}$ is a feasible solution but possibly not a global solution of* (12)-(17)*, with* (16) *replaced by*

$$\Phi\left[\sum_{t=1}^{T}\lambda_t\sum_{k=1}^{L}\alpha_j^k(t)\right] \le b_j(t), \qquad j=0,\ldots,k;\ 0\le\lambda_t\le 1;\ \sum_{t=1}^{T}\lambda_t=1,$$

and $b_j \ge \sum_{t=1}^{T}\lambda_t b_j(t)$.

6.1. A Numerical example

We provide here a very preliminary numerical example where the solution of the formulation (12)–(17) is considered, and the budget constraints (16) are replaced by the unique one (which aggregates possible different budgets at the epochs $1,\ldots,T$)

$$\sum_{k=1}^{L}\sum_{j=0}^{k}\sum_{t=1}^{T}\alpha_j^k(t) \le budget.$$

Since the multi-period formulation is indeed nonlinear and possibly nonconvex, we preferred to preliminarily investigate its solutions through a renowned nonlinear (NLP) solver from the literature. In particular, we adopted Knitro 13.2.0 from the NEOS serve. Nevertheless, we are aware that since the problem might be nonconvex (unless the last property is explicitly proved to hold for (12)–(17)), the choice of the starting point by the NLP solver may be crucial, to both outreach an accurate final solution and to reduce the overall computational burden. The parameters of the two instances we considered are summarized as follows (respectively):

$$\begin{cases} L=10 \\ a_k=1/L, \qquad k=1,\ldots,L, \\ P_+(t)=0.65, \\ T=6, \\ budget=20, \end{cases}$$

and

$$\begin{cases} L = 10, \\ a_k = 1/L, \qquad k = 1, \dots, L, \\ P_+(t) = 0.65, \\ T = 6, \\ budget = 100. \end{cases}$$

We will briefly show that due to the smaller budget (i.e. $20 < 100$) in the first scenario, it yields worse results (i.e. a smaller value of $P_+(T)$) with respect to the second one. Indeed, adopting the first set of parameters for our multi-period formulation, Knitro presolver was able to eliminate 403 constraints and 1 variable, so that the final formulation included 395 variables and 660 constraints. Moreover, for the probabilities $\{P_+(t)\}$ we obtained the following nonmonotone sequence (due to the nonlinearity of the multi-period model)

$P_+(1)$	0.65
$P_+(2)$	$9.21406e - 05$
$P_+(3)$	$1.51607e - 05$
$P_+(4)$	$2.37203e - 05$
$P_+(5)$	$2.06585e - 05$
$P_+(6)$	0.499547.

Conversely, adopting the second set of parameters for our multi-period formulation, Knitro presolver was again able to eliminate 403 constraints and 1 variable, yielding for the probabilities the values

$P_+(1)$	0.65
$P_+(2)$	0.140855
$P_+(3)$	0.0613564
$P_+(4)$	0.0521609
$P_+(5)$	0.0573217
$P_+(6)$	1.

Again we observe a nonmonotone behaviour for the sequence $\{P_+(t)\}$, due to nonlinearities of the multi-period model. For the sake of completeness we remark that the solution of the last two nonlinear instances by Knitro required no more than 10 seconds each.

7. Conclusions

We studied the problem of possibly enhancing a standard dynamics from sociophysics (namely the model (1)), using a mathematical programming perspective. We were interested about following a couple of lines of research. First we coupled the model (1) with a LP scheme, in order to possibly control the probability $P_+(t + 1)$ for a given probability $P_+(t)$, after leveraging the unknowns a_j^k in (2). The final value of these variables gives a measure of the effort which is required to steer the diffusion process of the opinion $+$ for the subgroups of cardinality k. The dynamics is strongly based on the idea of possibly maximizing $P_+(t + 1)$. A numerical experience is also reported in this regard, showing that the computational effort for solving LPs associated with this first proposal is definitely modest, even in case the solution of large instances is sought.

As a second task, we also extended the (one period) LP formulations to a more general multi-period formulation, being in this case necessary to compute the sequence $\{P_+(t + 1), P_+(t + 2), \dots, P_+(T)\}$ for a given value of $P(t)$ and for a given time horizon T. Some theoretical properties showing a connection between the solutions of both our proposals have been given, though much work yet requires to be carried on, including a wide numerical experience also in the multi-period case.

References

1. Backstrom, L.; Huttenlocher, D.; Kleinberg, J.; Lan, X. Group formation in large social networks: membership, growth, and evolution. In *KDD 2006: Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining* 2006,44–54.
2. Bass, F. M. A new product growth for model consumer durables. *Manage Sci* **1969**, *15(5)*, 215–227.
3. Ellero, A.; Fasano, G.; Favaretto, D. An application of Linear Programming to Sociophysics Models. In *Modeling and Analysis of Complex Systems and Processes - MACSPro'2020, October 22–24, 2020, Venice, Italy & Moscow, Russia*; Shapoval, A., V. Popov V., and I. Makarov, I.„Eds.; 2020, 23–33.
4. Ellero, A.; Fasano G.; Sorato, A. A modified Galam's model for word-of-mouth information exchange. *Physica A* **2009** *388(18)*, 3901–3910.
5. Ellero, A.; Fasano G.; Sorato, A. Stochastic Model for Agents Interaction with Opinion–Leaders. *Phys Rev E* **2013** *87(4)*, n. 042806.
6. Fourer, R. Modeling Languages versus Matrix Generators for Linear Programming. *ACM T Math Software* **1983** *9*, 143–183.
7. Fourer, R. *AMPL : a modeling language for mathematical programming*, 2nd ed.; Scientific Pr.: San Francisco, Calif, 1996.
8. Galam, S. Modelling rumors: the no plane pentagon french hoax case. *Physica A* **2003** *320*, 571–580.
9. Galam, S. Sociophysics: a review of Galam models. *Int J Mod Phys C* **2008** *19 (3)*, 409–440.
10. Galam, S. *Sociophysics. A Physicist?s Modeling of Psycho-Political Phenomena*; Springer Science+Business Media: New York, NY, 2012.
11. Galam, S.; Jacobs, F. The role of inflexible minorities in the breaking of democratic opinion dynamics. *Physica A* **2003** *381*, 366–376.
12. Moore, G. *Crossing the chasm*; Harper Collins: New York, NY, 1991.
13. Rogers, E. M. *Diffusion of Innovations*, 5th ed.; The Free Press: New York, NY, 2003.
14. Schelling, T. C. Dynamic models of segregation. *J Math Sociol* **1971**, *1*, 143–186.