Physica A 603 (2022) 127666

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

An example of social interaction: Spatial contagion effect in exams

José J. Arenas [a], Pedro Carpena [a,b,*]

[a] Departamento de Física Aplicada II, E.T.S.I. de Telecomunicación, Universidad de Málaga, 29071, Málaga, Spain
[b] Instituto Carlos I de Física Teórica y Computacional, Universidad de Málaga, 29071, Málaga, Spain

ARTICLE INFO

Article history:
Received 3 August 2021
Received in revised form 8 December 2021
Available online 3 June 2022

Keywords:
Social systems
Human interactions
Contagion effect

ABSTRACT

In the last decades the study of collective phenomena has produced a great interest in the field of Statistical Physics within the framework of Complex Systems, being a paradigmatic example flocking, the collective motion of self-propelled organisms. Such studies have been more recently extended to collective human behavior, where social interactions are important and concepts such as 'social force' have arisen. In this work we want to explore the possible existence of a 'social field' in a very controlled human environment: a classroom where the students take an exam. Since the students are seated in individual tables while working in their corresponding exams, the only possible interaction occurs when the students finish the exam and deliver it to the teacher. We conjecture that the existence of social interactions could lead to a contagion effect among the students, so that a given student who delivers the exam may influence another close student to do the same, and as a result the exams are not randomly delivered in the space. In this sense, each classroom can be seen as a complex system, where there exist interactions between the different elements, the students. To show the existence of this contagion effect, we use experimental data registered in 10 high-school classrooms during different exams, and for each student we record the exam delivery time and the spatial location of the student in the classroom while taking the exam. We use the distances between students who finish the exam consecutively and compare these distances with the random expectation in the corresponding classroom using Monte Carlo simulations. We observe a significant nonrandom behavior of the experimental data, and show the existence of a clustering effect in space, supporting the existence of a contagion effect as a consequence of an underlying 'social field'. Finally, to quantify this contagion effect, we propose a probabilistic distance-driven contagion model, according to which a given student who delivers the exam may influence another student closer than a given distance to do the same with certain contagion probability. By comparing the results of the model with the experimental data, we obtain a global contagion probability of around 1/6.

1. Introduction

In the last decades, collective phenomena have been the subject of Statistical Physics in the realm of Complex Systems. Probably, since the pioneering work of Vicsek and coworkers [1], flocking was one of the first examples considered: the

* Corresponding author at: Departamento de Física Aplicada II, E.T.S.I. de Telecomunicación, Universidad de Málaga, 29071, Málaga, Spain.
E-mail addresses: arenasferrer@hotmail.com, josearenas@uma.es (J.J. Arenas), pcarpena@ctima.uma.es, pjcarpena@uma.es (P. Carpena).

https://doi.org/10.1016/j.physa.2022.127666

collective and synchronized motion of self-propelled organisms (specially birds [2]), which has become an active area of research since then (see, for example, the review [3]). More recently, these studies have been extended to collective human behavior, leading to the field of Sociophysics, in which the collective behavior emerges from the interactions of individuals as elementary units in social structures (see for example the variety of examples in the review [4]). Within this context, the motion of pedestrians is probably one of the most active fields [5–9]. One of the models proposed and developed in some of these works to quantitatively describe pedestrian motion is the *social force model* [5,8,9], which contemplates the idea of social interaction. To do so, the social force model considers the individual response to the effect of the environment (other pedestrians and walls or borders).

In this work we want to introduce a new example of human interaction: the possible existence of a 'social field' in a classroom where a group of students take an exam. Our personal (subjective) observations from our teaching along the years suggest that the students do not finish the exam and deliver it to the teacher in a random manner in the space. On the contrary, we have the perception that very often (more often than expected by chance) students who deliver consecutively their exams are located in neighboring areas of the classroom. This phenomenon, if confirmed, would imply some degree of underlying self-organization, and would lead to the existence of a 'contagion effect' among students as a consequence of a social interaction. The contagion effect in social human behavior has been detected in many different contexts, including professional career mobility [10], and interpersonal influence concerning phenomena as diverse as obesity, smoking habits or cooperation [11]. In this sense, each classroom could be seen as a complex systems with interactions among students instead of a collection of non-interacting individuals. By using experimental data registered in several groups of students while taking different exams, the aim of this work is to show quantitatively that indeed there exists a spatial contagion effect among students.

We note that in an environment as controlled as a classroom with students taking an exam under the supervision of the teacher, the interaction between them is null meanwhile they work individually in the test. Therefore, the only possible influence from a given student to other(s) must take place only when the student finishes the exam and delivers it to the teacher. This event could act as a possible trigger (a 'perturbation') and make other students to do the same. As mentioned above, this contagion effect would be the consequence of an underlying social interaction, which may be produced by different reasons: friendship between some students, gregarious behavior, discomfort to deliver the exam in an isolated way thus becoming more conspicuous, etc.

We are aware that not all the students will be affected by this 'social field', and many of them will finish and deliver the exam freely. But if some of the students are indeed influenced by the social field, this should be detectable by analyzing the positions in the classroom of students who deliver the exam consecutively. Note that if no social interaction is present, given a student who delivers the exam, the next student to do so should be located randomly in the classroom, with the same probability of being close or far from the previous one. However, if a social spatial interaction is acting, one should expect the second student to be closer to the first than expected by chance. This kind of 'attraction' should produce some degree of spatial clustering for consecutive students. Such spatial clustering has been observed to spontaneously appear in human collectives provided there exist a visual connection between individuals [12]. Similar clustering behavior has also been observed in different types of physical systems, such as for example energy levels for long-range correlated disordered systems [13,14], and keywords in written texts and amino acids in proteins [15].

Our analysis is carried out by using experimental data registered in ten different classrooms. The students in each classroom were taking an exam (of different subjects), and we registered both the spatial location and the exam delivery time for each student in all classrooms. We include a full description of the experimental data and the protocol we followed to obtain them in Section 2.

In Section 3 we study the spatial behavior of students who deliver consecutively their exams. Each classroom is viewed as a $n \times m$ grid according to the geometric disposition of the tables in the classroom where the students were seated while taking the exam, with n the number of rows and m the number of columns. We use the sum of the distances between students who deliver consecutively their exams as an indicator of the existence of a social interaction, since we expect this quantity to be significantly smaller than expected by chance if the contagion effect exists. In each classroom, we obtain the experimental value of this measure and determine the corresponding statistical significance by means of Monte Carlo simulations. To do so, we generate a large number of random exam deliveries using the geometric configuration of the corresponding classroom.

Since the results show a clear nonrandom spatial behavior, in Section 4 we propose a probabilistic distance-driven contagion model between students, according to which a given student who delivers his/her exam may influence with certain probability other student located closer than a given distance. By comparing the results of the model and the experimental observations we can obtain a global short-distance contagion probability of around 1/6. Finally, we present our conclusions in Section 5.

2. Data

We use in our analyses the data corresponding to 10 exams (of several subjects) carried out in different classes totaling 264 students with ages in the range 12–16 years old. The students attended to two different high schools located at different cities in the south of Spain (Estepona and Córdoba). In each exam, the students were seated in individual tables arranged as a $n \times m$ grid with n the number of rows and m the number of columns, different in general for each classroom.

Table 1
Structure of the registered data in a given classroom. The number assigned to each student corresponds to the temporal order at which the student finished the exam and handed it to the teacher, and is linked to the corresponding delivery time and position in the classroom.

Student	Delivery time	Position
1	t_1	(x_1, y_1)
2	t_2	(x_2, y_2)
⋮	⋮	⋮
N	t_N	(x_N, y_N)

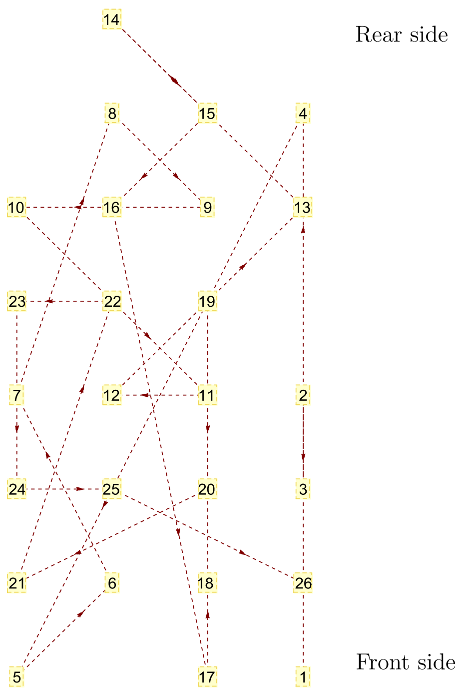

Fig. 1. The spatial configuration of a classroom with 26 13-year-old students while taking an exam of Physics and Chemistry. Each small square represents an occupied table, and the numbers within each square corresponds to the temporal order at which the student finished the exam and delivered it to the teacher. The dotted lines with arrows connect students who delivered the exam in consecutive order.

The specific protocol for any of the exams was the following: (i) The teacher calls the student to be seated in individual tables in alphabetical order, occupying the tables from the front side of the classroom to the rear part. (ii) Once all the available students are seated, the teacher communicates the prescribed time for finishing the exam, and the exam begins. (iii) If some students do not deliver their exam within the allowed time, the teacher makes a final call and the remaining students deliver their exams within a very short period (typically, 2–3 min). We note that not all the tables were occupied since empty places could appear if some students do not attend to the exam. The number of students also changes for different classrooms. In order to record the data, we followed the same procedure in all cases: we provided the teacher with a template where he/she annotated sequentially the time (in minutes) at which each student finished the exam and delivered it to the teacher, as well as the corresponding position in the $n \times m$ grid where the student was seated while taking the exam, with the origin $(1, 1)$ located at the left corner of the front row. The names of the students were not registered, to preserve the anonymity. In addition, the students were not informed of the experiment in order to prevent biased results, so that the students were completely free to deliver their exams when they wanted to.

In this way, once a particular exam finished in a classroom with N students, each student was identified with a number j $(j = 1, 2 \ldots, N)$ corresponding to the temporal order at which the student finished the exam and delivered it to the teacher, being 1 the first and N the last student. For any j, we also had the information of the corresponding delivery time t_j, with $t_j \leq t_{j+1}$. In addition, each student j was also linked to his/her corresponding position (x_j, y_j) in the classroom, where x_j and y_j are integers in the range $1, \ldots, n$ and $1, \ldots, m$ respectively. The final data recorded in a given classroom follows the structure given in Table 1.

An example of one of the considered classrooms is shown in Fig. 1, where a group with 26 13-years-old students is depicted. The students are placed in a 7×4 grid, with some empty positions. In this figure, each small square represents an occupied table in the grid, and the number j within each square represents the temporal ordering at which the corresponding student delivered his/her exam to the teacher. The dotted lines with arrows connect students who delivered the exam in consecutive order, i.e., students j and $j+1$ with $j = 1, \ldots, N-1$.

As we have mentioned in the Introduction, we conjecture that there exists a social interaction between students which ultimately produces a contagion effect leading to a non-random delivery of exams. We expect this interaction or 'social field' to act at short distances, i.e. a given student who finishes the exam and hands it to the teacher can influence neighboring students (close in space) to deliver their exams too. We analyze if this is the case in the next section.

3. Spatial analysis

In this section, we want to study a possible spatial interaction between the students. In particular, if this 'contagion effect' exists, given a student who delivers his/her exam, with a higher probability than expected by chance the following exam delivered to the teacher should correspond to a student located at short distance from the previous one. In order to quantitatively determine if this is the case, we propose the following approach: Let us consider one of the ten classrooms with registered data, and let us assume that there are N students in the classroom. First we calculate the spatial distances d_j between students j and $j+1$, i.e. students who deliver their exams consecutively. To show that the results are independent of the particular details of the distance considered, we use two different distances to quantify the proximity between these students:

$$\begin{aligned} d_j^{(1)} &= \sqrt{(x_{j+1}-x_j)^2 + (y_{j+1}-y_j)^2} \\ d_j^{(2)} &= \max(|x_{j+1}-x_j|, |y_{j+1}-y_j|) \end{aligned} \quad (1)$$

and in both cases, $j = 1, \ldots, N-1$. With these definitions, $d^{(1)}$ corresponds to the classical Euclidean distance, and $d^{(2)}$ measures the separation as the number of rows or columns between the students. In both cases, the unit of distance is the separation between rows or columns, assumed to be the same. Second, we note that each individual distance value $d_j^{(i)}$ $(i = 1, 2)$ accounts for a possible local interaction between two students, j and $j+1$. However, to measure the global contagion effect in the whole classroom, we propose to obtain the experimental 'walk' W_{exp} for the classroom as:

$$W_{exp}^{(i)} = \sum_{j=1}^{N-1} d_j^{(i)} \quad (2)$$

where i can be either 1 or 2, depending of the distance considered according to Eq. (1). For a given classroom, as for example the one shown in Fig. 1, W_{exp} measures the total length separating consecutive students in the whole classroom. We expect that if a social interaction is present the value of W_{exp} should be small as compared to the expectation of a random delivery of exams in the classroom. Note that if a given student influences other close students, an overrepresentation of small d_j values are expected as compared to pure randomness, and this effect should globally appear in W_{exp}.

To check that this is the case, we use Monte Carlo simulations to produce exams delivered randomly in each classroom. To do so, we proceed as follows: for a given class with N students, each student j is linked to his/her position (x_j, y_j). The real delivery order of the exam is $1, 2, \ldots, N$. Therefore, a random delivery of exams in that classroom is given by a permutation of the set of integers $(1, 2, \ldots, N)$. Since each integer (student) is linked to a position in the grid, we can calculate the $N-1$ distances between consecutive students according to the corresponding permutation and, by using an expression similar to Eq. (2), the corresponding random walk $W_{ran}^{(i)}$. Finally, we generate a large number of permutations n_p in each classroom to determine the statistical behavior of $W_{ran}^{(i)}$.

An example of the results of the simulation is shown in Fig. 2, where we plot the probability density of W_{ran} obtained numerically using the above algorithm for the classroom depicted in Fig. 1 by generating $n_p = 10^6$ permutations. We show two probability densities, $p(W_{ran}^{(1)})$ (Fig. 2a) and $p(W_{ran}^{(2)})$ (Fig. 2b), corresponding to the use in Eq. (2) of the distances $d_j^{(1)}$ and $d_j^{(2)}$ respectively. We also indicate in both panels the mean, median and standard deviation for each distribution. The experimental walks obtained in the classroom are given by $W_{exp}^{(1)} = 59.98$ and $W_{exp}^{(2)} = 54$, and they are also marked in Figs. 2a and 2b. In both cases, the experimental $W_{exp}^{(i)}$ values fall well into the left tail of the corresponding distribution, thus indicating that there is an overrepresentation of short distances between consecutive students which can be understood as a signature of social interaction via contagion effect. Indeed, with the results of the simulation we can calculate for each case the probability of obtaining an equal or smaller value (p-value) than the experimental one, and we find respectively $\text{prob}(W_{ran}^{(1)} \leq W_{exp}^{(1)} = 59.982) = 0.009294$ and $\text{prob}(W_{ran}^{(2)} \leq W_{exp}^{(2)} = 54) = 0.011391$. Both results are quite similar, indicating first that they do not depend on the specific details of the distance considered (provided it is reasonably defined), and second, and more important, the existence of a contagion effect or 'social field' with a probability of around 99% in the corresponding classroom.

We have carried out similar analyses to the one described above in the rest of the classrooms with available data. In all cases, the resulting probability densities are very similar in shape to the ones shown in Fig. 2: they exhibit a Gaussian-like profile, with means and medians almost identical. The statistical results for the 10 classrooms are summarized in Table 2.

Table 2
Summary of the output of the Monte Carlo simulations in the ten classrooms, with $n_p = 10^6$ permutations in each one. We include the number N of students in each classroom, the experimental values of $W_{exp}^{(1)}$ and $W_{exp}^{(2)}$, the corresponding p-values obtained numerically from the probability densities $p(W_{ran}^{(1)})$ and $p(W_{ran}^{(2)})$, and the mean of both distributions, $\langle W_{ran}^{(1)}\rangle$ and $\langle W_{ran}^{(2)}\rangle$. The classrooms are ordered according to the age of the students (younger—top, older—bottom).

N	$W_{exp}^{(1)}$	$\langle W_{ran}^{(1)}\rangle$	p-value	$W_{exp}^{(2)}$	$\langle W_{ran}^{(2)}\rangle$	p-value
30	81.412	84.618	0.295	72	75.536	0.227
24	60.056	65.711	0.157	54	59.667	0.169
26	59.982	74.914	0.009	54	68.460	0.011
31	85.176	89.762	0.238	76	80.642	0.245
30	86.247	87.523	0.416	78	78.669	0.484
21	53.678	52.038	0.638	48	46.858	0.647
26	62.376	72.130	0.055	57	65.690	0.081
24	50.327	64.985	0.003	48	58.664	0.022
22	54.708	58.306	0.249	50	53.446	0.279
30	83.319	84.615	0.410	73	75.533	0.350

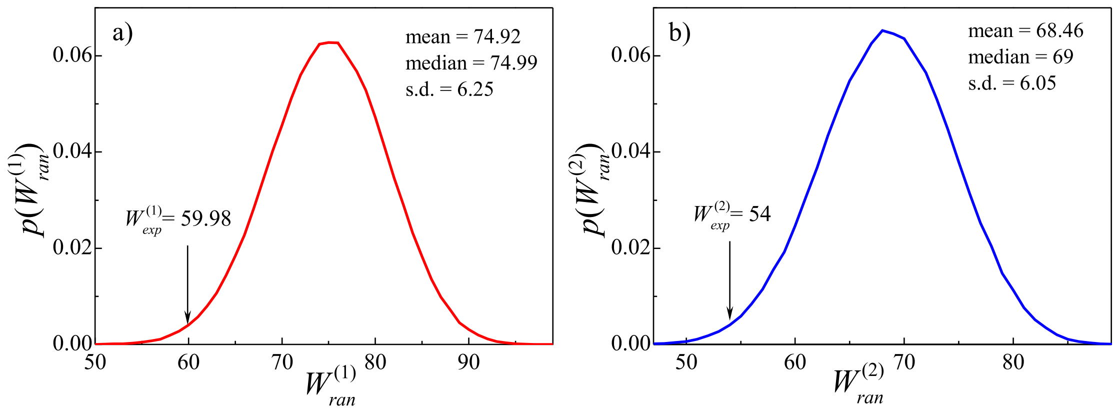

Fig. 2. Probability densities of the walks $W_{\mathrm{ran}}^{(1)}$ (panel (a)) and $W_{\mathrm{ran}}^{(2)}$ (panel (b)). Both have been obtained by generating 5×10^6 random permutations of the ordering of the students in the classroom shown in Fig. 1. In both panels, we also indicate with an arrow the real experimental values $W_{exp}^{(1)}$ and $W_{exp}^{(2)}$. We include in each case the values of the mean, median and standard deviation.

We note that in 9 out of 10 classrooms the experimental walks $W_{exp}^{(1)}$ and $W_{exp}^{(2)}$ are smaller than the means (and medians) of the corresponding stochastic variables $W_{ran}^{(1)}$ and $W_{ran}^{(2)}$, thus suggesting that the contagion effect, although stronger in some classrooms than in others, is a general phenomenon, not depending on the particular classroom considered.

In order to analyze a global statistical significance of the results, we can also define a total walk $W_{\mathrm{tot,exp}}$ given by the sum of the experimental walks obtained in each classroom:

$$W_{tot,exp}^{(i)} = \sum_{k=1}^{n_c} W_{k,exp}^{(i)} \tag{3}$$

where i can be 1 or 2, n_c is the number of classrooms considered ($n_c = 10$ in our case), and the index k runs over all the classrooms. As we have calculated already $W_{exp}^{(1)}$ and $W_{exp}^{(2)}$ in each classroom (Table 2), we finally obtain $W_{tot,exp}^{(1)} = 677.281$ and $W_{tot,exp}^{(2)} = 610$. These empirical values can be tested statistically by defining the stochastic variable 'total random walk', $W_{tot,ran}$, as:

$$W_{tot,ran}^{(i)} = \sum_{k=1}^{n_c} W_{k,ran}^{(i)} \tag{4}$$

where $W_{k,\mathrm{ran}}^{(i)}$ stands for a random walk in the k−th classroom. By generating a large number of student permutations n_p in each classroom ($n_p = 10^6$), we obtain numerically the probability distributions of $W_{tot,ran}^{(i)}$, with $i = 1, 2$. Both probability densities are shown, respectively, in Figs. 3(a) and 3(b). We include in each panel of Fig. 3 the experimental values $W_{tot}^{(1)}$ and $W_{tot}^{(2)}$ (marked with arrows), which are located at the farthest extreme of the left tail of the distribution in both cases, at a distance of about 3 standard deviations of the corresponding mean, reinforcing the idea of the existence of a spatial contagion effect. Indeed, using the two numerically determined probability distributions, we can obtain

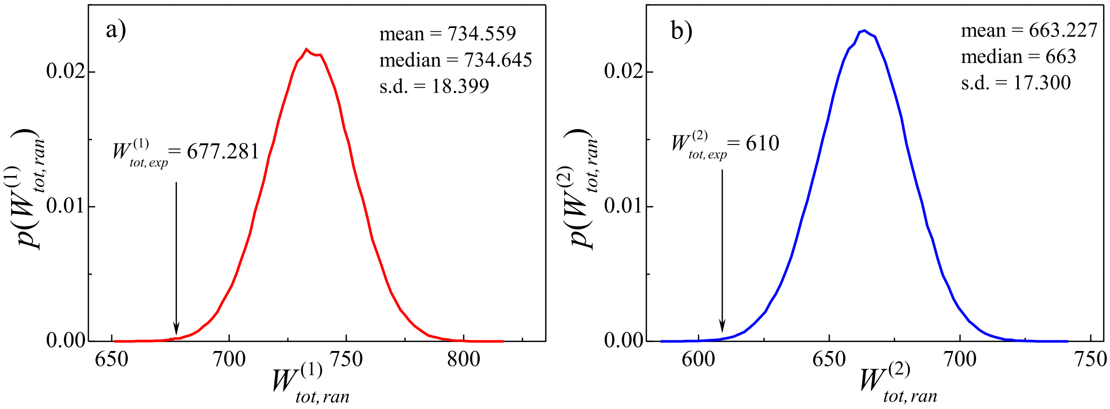

Fig. 3. Probability densities obtained by Monte Carlo simulations ($n_p = 10^6$) of the stochastic variables $W^{(1)}_{tot,ran}$ (panel (a)) and $W^{(2)}_{tot,ran}$ (panel (b)). We also include in each panel the mean, median and standard deviation of the corresponding distribution. The respective experimental values $W^{(1)}_{tot,exp}$ and $W^{(2)}_{tot,exp}$ are indicated with arrows.

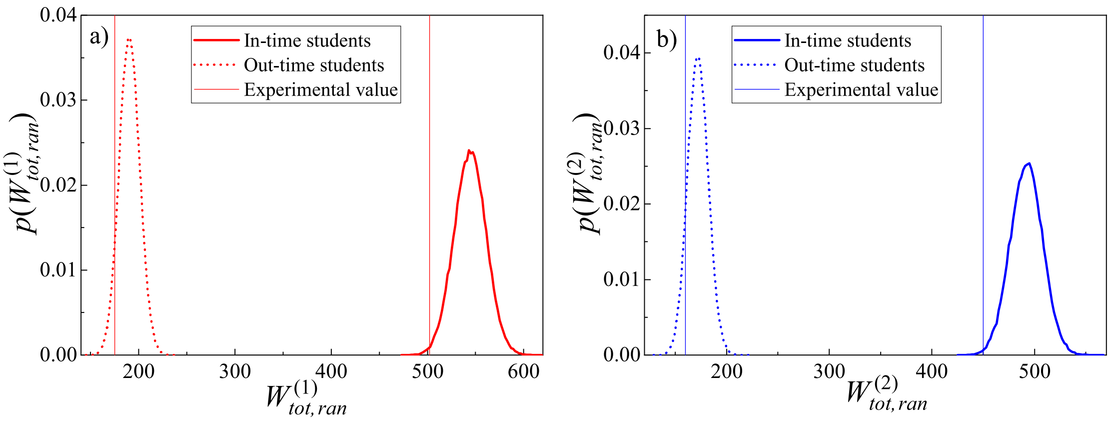

Fig. 4. Probability densities obtained by Monte Carlo simulations ($n_p = 10^6$) of the stochastic variables $W^{(1)}_{tot,ran}$ (panel (a)) and $W^{(2)}_{tot,ran}$ (panel (b)), but by considering separately the students who deliver the exam within the allowed time, and those who deliver their exams out of time. The vertical lines correspond to the experimental value $W^{(j)}_{tot}$ obtained for each group.

the statistical significance of $W^{(1)}_{tot,exp}$ and $W^{(2)}_{tot,exp}$, and we get prob($W^{(1)}_{tot,ran} \leq W^{(1)}_{tot,exp} = 677.281$) $= 0.00103$ and prob($W^{(2)}_{tot,ran} \leq W^{(2)}_{tot,exp} = 610$) $= 0.00123$. Remarkably, both p-values are very similar, of around 10^{-3}, so that the detection of the spatial contagion effect does not depend on the particular distance considered.

This global small p-value supports the existence of a short-range spatial interaction between consecutive students (contagion effect), which seems to be a general property, not specific of a single classroom: in general, the proximity between students produces a high probability of consecutive delivery of exams (much higher than expected by chance). We remark that this result has been obtained by considering the distance between consecutive students, i.e., the possible influence of a student on the next one in the temporal order.

We note that the global result shown in Fig. 3 has been obtained by considering the positions in the classroom of all the students, both the ones who deliver the exam before the final time fixed by the teacher and those who deliver the exam after this final time, forced by the teacher to do so in a very limited time interval. Since we know the positions in the classroom and the temporal ordering for all the students, we can obtain separately the experimental value of $W^{(j)}_{tot,exp}$ for in-time students and for out of time students, and compare these experimental values with the results of Monte Carlo simulations similar to the ones described above ($n_p = 10^6$ permutations in each classroom) but carried out separately for both sets of students. Concerning the experimental values, we obtain $W^{(1)}_{tot,exp} = 502.115$ and $W^{(2)}_{tot,exp} = 450$ for the set of in-time students in all classrooms, while for the set of out of time students, $W^{(1)}_{tot,exp} = 175.17$ and $W^{(2)}_{tot,exp} = 160$. The comparison of these experimental values with the results of the corresponding Monte Carlo simulations are shown in Fig. 4. We obtain that prob($W^{(1)}_{tot,ran} \leq 502.115$) $= 0.00466$ and prob($W^{(2)}_{tot,ran} \leq 450$) $= 0.00396$ for in-time students, while for out of time students we get prob($W^{(1)}_{tot,ran} \leq 175.17$) $= 0.0739$ and prob($W^{(2)}_{tot,ran} \leq 160$) $= 0.130$. According to these values, and although the contagion effect is clearly stronger (more significant) for the set of students who deliver their exams on time than for out of time students, the contagion effect is also present in this second group. This is the reason why when both sets are considered together and the sample becomes larger, the corresponding p-value decreases and the global result becomes more significant (see Fig. 3). For this reason, when we develop a contagion model in Section 4 (see below) we do not separate between both sets of students in each classroom.

We can also investigate the possible interaction between non-consecutive students by calculating experimental values similar to (2) for individual classrooms and to (3) for a global behavior, but considering the corresponding distance $d_j^{(i)}$ in both expressions as the distance between second temporal neighbors, i.e., between students j and $j+2$. The statistical significance of these experimental values can be then checked by computing the random expectation via Monte Carlo simulations, in order to obtain the probability distribution of $W_{ran}^{(i)}$ in each classroom, and of the global variable $W_{tot,ran}^{(i)}$ considering in both cases the distances between second temporal neighbors. However, in this case the results indicate that there is no significant deviation of the experimental values from the random expectation obtained from Monte Carlo simulations. As an example, using the Euclidean distance $d_j^{(1)}$ between second temporal neighbors in Eq. (3), we obtain the global result $W_{tot,exp}^{(1)} = 712.635$. By means of Monte Carlo simulations ($n_p = 10^6$ permutations in each classroom), and similarly to what we showed in Fig. 3 but for second temporal neighbors, we have calculated the probability distribution $p(W_{tot,ran}^{(1)})$ from where we can obtain $\text{prob}(W_{tot,ran}^{(1)} \leq W_{tot,exp}^{(1)} = 712.635) = 0.642$. This non-significant p-value, obtained using distances between second temporal neighbors, is in clear contrast with the significant p-value ($\sim 10^{-3}$) obtained for consecutive students. These two antagonistic results suggest that the contagion effect only acts between consecutive students, and the influence of a student on the nth temporal neighbor can be disregarded for $n > 1$. For this reason, in the next section we develop a probabilistic contagion model that works only for consecutive students, and that is able to reproduce the experimental observations and quantify the contagion probability between such students.

4. Contagion model

The results of the previous section indicate the existence of a spatial clustering for students who deliver consecutively their exams, since in general we observe a higher proximity between such students than expected by chance. This effect must be produced by some spatial interaction (contagion) which should act at short distances. Based on this idea, we propose a probabilistic contagion model depending on the distance between students. In this way, the model has two input parameters: the contagion probability p and the maximum distance $d_{\max}$ at which the contagion acts. The model works as follows:

(1) Let us consider a classroom with N students. The first student to deliver the exam, s_1, is picked at random from the N students present in the classroom.

(2) After j students have delivered their exams, the last student to do so is s_j. The next student to deliver the exam (s_{j+1}) is chosen in the following way: first, we calculate the distances from s_j to the $N-j$ students remaining in the classroom, and determine the m students for which such distance is less or equal than $d_{\max}$. Then, with probability p we choose at random s_{j+1} among the m students closer than $d_{\max}$, and with probability $1-p$, we choose s_{j+1} at random from all the $N-j$ remaining students. In this way, p quantifies the contagion probability below $d_{\max}$. An example illustrating how the contagion model works is shown in Fig. 5.

We can formalize mathematically the model in the following way: when j students have delivered their exams, the last student to do so is s_j. At this moment, there are $N-j$ students remaining in the classroom, so that m students are closer to s_j than $d_{\max}$ and $N-j-m$ are farther than $d_{\max}$. Then, the probability for any of the m close students of being the next one to deliver the exam is:

$$\text{Prob}_m = \begin{cases} \dfrac{p}{m} + \dfrac{1-p}{N-j} & \text{if} \quad m \neq 0 \\ 0 & \text{if} \quad m = 0 \end{cases} \tag{5}$$

where we have considered the case $m = 0$, i.e. the situation in which there are no students closer to s_j than $d_{\max}$. In the more interesting case $m \neq 0$, the first term on the RHS of the equation is the contagion term, so that the contagion probability p is shared by the m close students. The second term accounts for the random expectation, which is the same for all the $N-j$ students that remain in the classroom. Similarly, the probability for any of the $N-j-m$ students farther than $d_{\max}$ is then:

$$\text{Prob}_{N-j-m} = \begin{cases} \dfrac{1-p}{N-j} & \text{if} \quad m \neq 0 \\ \dfrac{1}{N-j} & \text{if} \quad m = 0 \end{cases} \tag{6}$$

Obviously, the normalization condition is fulfilled since for both cases $m \neq 0$ and $m = 0$ it straightforward to check that:

$$m \cdot \text{Prob}_m + (N-j-m) \cdot \text{Prob}_{N-j-m} = 1 \tag{7}$$

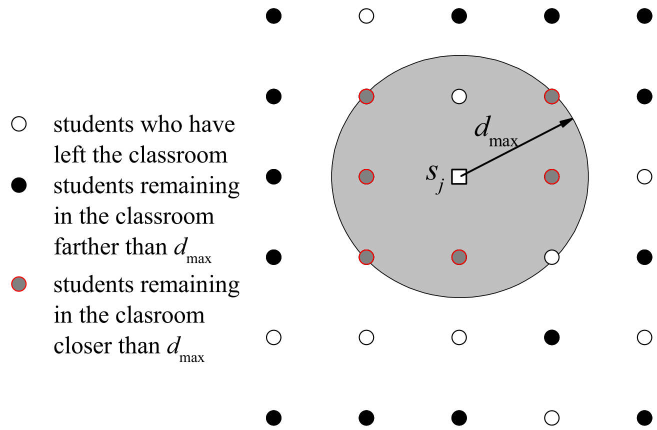

Fig. 5. A schematic representation of the contagion model. Open circles correspond to empty tables (students who have delivered the exam and left the classroom), while solid circles represent students still working in their exams. The last student to deliver the exam is s_j (open square). This student may influence other students still remaining in the classroom closer than d_{max} (solid gray circles). In this case, the number of such students is $m = 6$. The probability of being the next one to deliver for any of these m close students is given in Eq. (5). For any of the students that remain in the classroom which are further than d_{max} (black circles), the corresponding probability is given in Eq. (6). In this figure, $d_{max} = \sqrt{2}$ if $d_j^{(1)}$ is used, while $d_{max} = 1$ in the $d_j^{(2)}$ case.

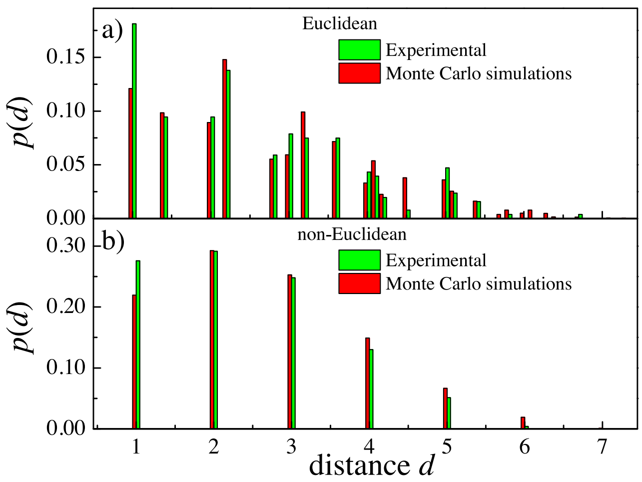

Fig. 6. Probability densities $p(d)$ of the distances between students who deliver their exams in consecutive order observed experimentally and obtained numerically using Monte Carlo simulations by grouping all the classrooms. Panels (a) and (b) correspond, respectively, to the use of the distances $d^{(1)}$ and $d^{(2)}$ of Eq. (1).

Once an iteration of the model is run in a given classroom, we obtain a set of ordered students $\{s_1, s_2, \ldots, s_N\}$, from which a particular value of the stochastic variable $W(p, d_{\max})$ can be obtained as

$$W^{(i)}(p, d_{\max}) = \sum_{j=1}^{N-1} d_j^{(i)}(p, d_{\max}) \tag{8}$$

with $d_j^{(i)}(p, d_{\max})$ the distance between students s_j and s_{j+1} obtained running the model using parameters p and $d_{\max}$, and $i = 1, 2$ depending on the distance considered (see Eq. (1)).

This contagion model depends on two parameters, p and d_{max}. However, we expect the student interaction to be short-ranged, therefore restricting d_{max} to small values. Indeed, we can confirm this short-range interaction and estimate *a priori* the value of d_{max} by computing the probability density $p(d)$ of the distances between consecutive students observed experimentally and the obtained from Monte Carlo simulations grouping all the classrooms to improve the statistics. The results for $p(d)$ are shown in Fig. 4, where we plot the normalized $p(d)$ in both cases obtained by using the Euclidean (panel (a)) and non-Euclidean (panel (b)) distances $d^{(1)}$ and $d^{(2)}$ in Eq. (1).

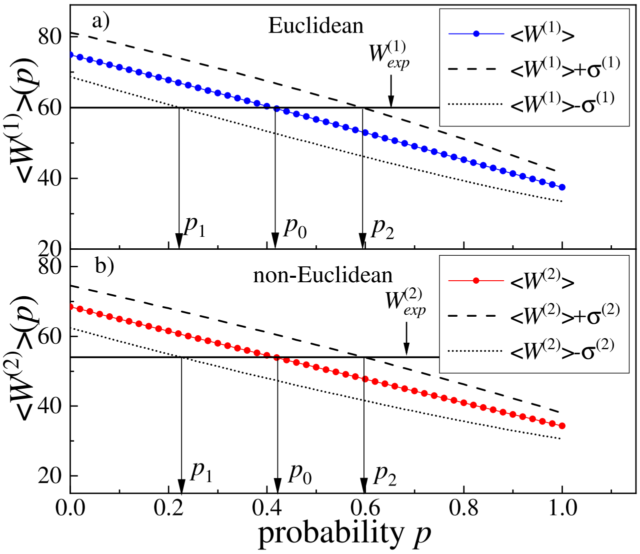

Fig. 7. (a) Expected value of $W(p)$ as a function of the contagion probability p obtained for the same classroom shown in Fig. 1 using the Euclidean distance $d_j^{(1)}$ of Eq. (1). The dashed and dotted lines represent, respectively, the expected value $\pm$ the standard deviation σ, which also depends on p. (b) The same as in panel (a), but using the non-Euclidean distance of Eq. (1). In both panels (a) and (b) the horizontal lines correspond to the experimental W_{exp} values obtained in that classroom, and the results have been obtained by running the contagion model $n_{mod} = 10^6$ times for each p value. See the text for a description of p_0, p_1 and p_2.

Remarkably, the most relevant characteristic shown in Fig. 6 is a clear increase of the probability of the experimental $d = 1$ case as compared to the random expectation. This increase is around 50% of the Monte Carlo result when considering the Euclidean distance, and around 25% of the random expectation for the non-Euclidean case. This result confirms the short-range character of the contagion effect, that seems to work only for nearest neighbors ($d = 1$).

The results shown in Fig. 6 suggest to consider $d_{\max} = 1$ as an appropriate input value for the contagion model, since this is the most favored distance when comparing the experimental results and the random expectation obtained via Monte Carlo simulations for both the $d^{(1)}$ and $d^{(2)}$ distances, so in the following we study the behavior of $W(p, d_{\max} = 1) \equiv W(p)$. In this way, for each individual classroom, we can run the contagion model a large number n_{mod} of times for any value of p to obtain numerically the probability distribution of $W(p)$, and calculate the expected value $\langle W\rangle(p)$. This expected value can be then compared to the experimental W_{exp} value obtained in the same classroom, so that the contagion probability in the classroom can be estimated as the p value for which the equality

$$W^{(i)}_{exp} = \langle W^{(i)}\rangle(p) \tag{9}$$

holds, and where $i = 1$ or 2, depending on the distance considered according to Eq. (1). From now on, we term p_0 the value of p satisfying (9), and corresponds to the estimated contagion probability in the corresponding classroom. In order to estimate an error interval for p_0, and since we can also determine the standard deviation $\sigma(p)$ of $W(p)$, we can consider the solutions of the equation:

$$W^{(i)}_{exp} = \langle W^{(i)}\rangle(p) \pm \sigma^{(i)}(p) \tag{10}$$

with $i = 1$ or 2. We term p_1 and p_2 the p values that are solutions of (10) for the '-' and '+ ' signs, respectively. In this way, p_1 and p_2 correspond to the limiting contagion probabilities for which W_{exp} lies within the interval $\langle W\rangle(p) \pm \sigma(p)$. An example of the behavior of $\langle W\rangle$ as a function of the contagion probability p and of the obtention of p_0, p_1 and p_2 is shown in Fig. 7, where we have used the results for the same classroom considered in Figs. 1 and 2. For each p value, we have run $n_{mod} = 10^6$ times the contagion model in the classroom to obtain $\langle W\rangle(p)$ and $\sigma(p)$. In this example, we have obtained an expected contagion probability $p_0 = 0.411$ with an error interval $[p_1, p_2] = [0.222, 0.593]$ for the Euclidean $d^{(1)}$ distance, and $p_0 = 0.416$ and $[p_1, p_2] = [0.225, 0.596]$ for the non-Euclidean $d^{(2)}$ case.

Proceeding similarly as we have done in this latter example, we have run the contagion model $n_{mod} = 10^6$ times in all the available classrooms, and obtained in each case the expected contagion probability p_0 and the corresponding error interval $[p_1, p_2]$, and the results are summarized in Table 3.

Several conclusions can be drawn from the results in Table 3: First, we observe quite similar values for p_0, p_1 and p_2 for the two $d^{(1)}$ and $d^{(2)}$ distances, thus indicating that the contagion effect can be detected independently of the used distance, provided that the distances are able to quantify reasonably the proximity among students. Second, the expected contagion probability p_0 depends on the particular classroom considered. This is not surprising, since the different classrooms are

Table 3
Expected contagion probability p_0 and the error interval $[p_1, p_2]$ for the ten classrooms shown in Table 2 obtained by using in the contagion model the Euclidean distance $d^{(1)}$ and the non-Euclidean distance $d^{(2)}$. In both cases, we have run the model $n_{mod} = 10^6$ times in each classroom for any value of p to solve numerically Eqs. (9) and (10), similarly to the example shown in Fig. 7.

Distance $d^{(1)}$		Distance $d^{(2)}$	
p_0	$[p_1, p_2]$	p_0	$[p_1, p_2]$
0.0748	[0,0.234]	0.0913	[0,0.249]
0.184	[0.00207,0.386]	0.188	[0.00965,0.381]
0.411	[0.222,0.593]	0.416	[0.225,0.597]
0.102	[0,0.262]	0.109	[0,0.266]
0.0292	[0,0.191]	0.0160	[0,0.173]
0	[0,0.140]	0	[0,0.161]
0.267	[0.0939,0.448]	0.257	[0.0789,0.442]
0.459	[0.272,0.633]	0.372	[0.182,0.558]
0.128	[0.0.335]	0.132	[0,0.342]
0.0304	[0,0.188]	0.0657	[0,0.221]

Table 4
Estimated global contagion probability p_g and the corresponding error interval $[p_{g1}, p_{g2}]$ obtained by using in the contagion model the Euclidean distance $d^{(1)}$ and the non-Euclidean distance $d^{(2)}$.

Distance	p_g	$[p_{g1}, p_{g2}]$
$d^{(1)}$	0.159	[0.105, 0.215]
$d^{(2)}$	0.158	[0.103, 0.215]

not homogeneous: the possible social links between students change in different classrooms, as well as the number of individuals with gregarious behavior, etc. And third, the expected contagion probability is directly related to the statistical significance of the experimental W_{exp} value for each classroom (see Table 2). Note that a high p_0 value implies a strong contagion between close students, which in turns indicates a low W_{exp} value (much smaller than expected by chance) and therefore a very small p-value and high statistical significance. In this sense, the only case where the estimated p_0 value is 0 corresponds to the only classroom where the corresponding W_{exp} value is larger than the $\langle W_{ran}\rangle$ value (see Table 2).

Similarly to what we did in Section 3, we can try to obtain a global contagion probability by applying the contagion model with the same probability p to all the classrooms, and calculate the global stochastic variable $W_{tot}(p)$ as

$$W_{tot}^{(i)}(p) = \sum_{k=1}^{n_c} W_k^{(i)}(p) \tag{11}$$

where $W_k^{(i)}(p)$ is the output obtained after running once the contagion model in the kth classroom using probability p, and n_c is the total number of classrooms ($n_c = 10$), with $i = 1, 2$. By running the contagion model a great number of times n_{mod} in each classroom, we can obtain a large number of values of $W_{tot}^{(i)}(p)$, from where the corresponding expected value $\langle W_{tot}^{(i)}\rangle(p)$ and standard deviation $\sigma_{tot}^{(i)}(p)$ can be obtained.

In this way, we can compare $\langle W_{tot}\rangle(p)$ and the global experimental result $W_{tot,exp}$ of Eq. (3), and estimate the global contagion probability p_g as the value of p for which the equation

$$W_{tot,exp}^{(i)} = \langle W_{tot}^{(i)}\rangle(p) \tag{12}$$

holds, with $i = 1, 2$ depending on the distance considered. Similarly to what we did for the results in individual classrooms, the error interval of p_g, which we term $[p_{g1}, p_{g2}]$, can be obtained as the values of p which are solutions of the equation

$$W_{tot,exp}^{(i)} = \langle W_{tot}^{(i)}\rangle(p) \pm \sigma_{tot}^{(i)}(p) \tag{13}$$

with p_{g1} the solution with the '-' sign, and p_{g2} with the '+ ' one. The behavior of $\langle W_{tot}^{(i)}\rangle(p)$ as a function of p is shown in Fig. 8, where we also show the graphical solutions of p_g and of the error interval $[p_{g1}, p_{g2}]$ for both the Euclidean (panel (a)) and non-Euclidean (panel (b)) distances. The numerical solutions of Eqs. (12) and (13) are summarized in Table 4.

According to these results, the estimated global contagion probability p_g and the error interval $[p_{g1}, p_{g2}]$ are almost identical for both distances, $p_g \simeq 1/6$, thus supporting the robustness of the result. This coincidence is also in agreement with the similar p-values ($\sim 10^{-3}$) we obtained in Section 3 for the experimental values $W_{tot,exp}^{(1)}$ and $W_{tot,exp}^{(2)}$ when tested against the corresponding random expectations (see Fig. 3).

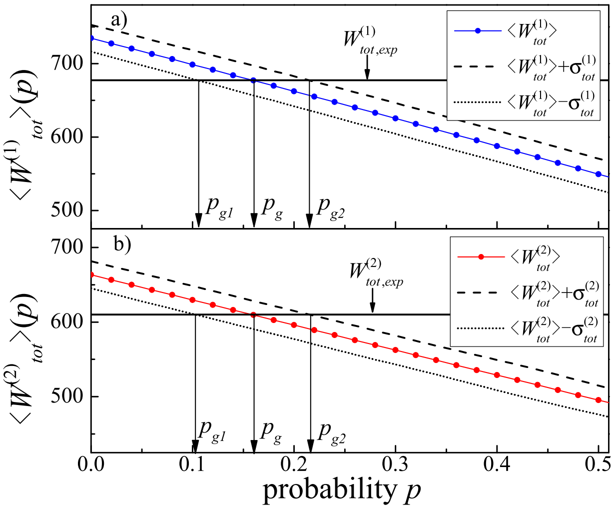

Fig. 8. Expected value $\langle W^{(i)}_{tot}\rangle(p)$ as a function of the contagion probability p. The dashed and dotted lines correspond to $\langle W^{(i)}_{tot}\rangle(p) + \sigma^{(i)}_{tot}(p)$ and $\langle W^{(i)}_{tot}\rangle(p) - \sigma^{(i)}_{tot}(p)$ respectively, where the $i = 1$ case (Euclidean) is shown in panel (a), and panel (b) corresponds to the $i = 2$ (non-Euclidean) case. We also indicate graphically how to estimate p_g, p_{g1} and p_{g2}, which are the solutions of Eqs. (12) and (13).

5. Conclusions

In this work we have studied an example of human social interaction in a very controlled environment: a classroom where the students take an exam. The students are seated in individual tables arranged geometrically in a $n \times m$ grid, and work individually in their exams. The only possible interaction between them can only occur when a student finishes the exam and delivers it to the teacher. If no social interaction is present, this event should happen randomly in space. However, we have shown that the consecutive exam deliveries of students do not happen randomly in space. We have obtained this result by comparing experimental data of the positions in the classroom of students who deliver the exam in consecutive order with the corresponding random expectations obtained by means of Monte Carlo simulations. We observe a general and clear clustering behavior when studying separately individual classrooms, which is also very statistically significant when all the classrooms are considered globally, thus supporting the existence of a spatial contagion effect between consecutive students as a consequence of the social interactions. In this sense, each classroom can be seen as a complex system with interactions between students instead of a collection of non-interacting individuals. To quantify this contagion effect, we have proposed a probabilistic distance-driven contagion model between students, according to which one student who delivers the exam may influence another student closer than a given distance to do the same with certain contagion probability. By comparing the results of the model with the experimental results, we can obtain an expected contagion probability both for each individual classroom and for all the classrooms globally considered. Although the contagion probability depends on the classroom considered, we can estimate a global contagion probability of around 1/6. This spatial contagion effect could also appear in other social human activities where the interaction between subjects occurs via single events (similar to an exam delivery), such as for example raised hand votation procedures in assemblies.

CRediT authorship contribution statement

José J. Arenas: Conceptualization, Software, Formal analysis, Writing (editing). **Pedro Carpena:** Methodology, Software, Formal analysis, Writing – original draft.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments

We wish to thank to all the teachers who recorded the experimental data for their help, and to J. M. Caballero for kindly reproducing some of the results. We also acknowledge the Spanish Junta de Andalucía (grants no. PIV-011/16 and FQM-362) and the Spanish Ministerio de Ciencia e Innovación (Grant No. PID2020-116711GB-I00) for financial support. Funding for open access charge: Universidad de Málaga/CBUA.

References

[1] T. Vicsek, A. Czirok, E. Ben-Jacob, I. Cohen, O. Shochet, Phys. Rev. Lett. 75 (1995) 1226.
[2] J.R. Usherwood, M. Stavrou, J.C. Lowe. K. Roskilly, A.M. Wilson, Nature 474 (2011) 494.
[3] T. Vicsek, A. Zafeiris, Phys. Rep. 517 (2012) 71.
[4] C. Castellano, S. Fortunato, V. Loreto, Rev. Modern Phys. 81 (2009) 591.
[5] D. Helbing, P. Molnar, Phys. Rev. E 51 (1995) 4282.
[6] I. Karamouzas, B. Skinner, S.J. Guy, Phys. Rev. Lett. 113 (2014) 238701.
[7] J.C. Moreno, M.L. Rubio Puzzo, W. Paul, Phys. Rev. E 102 (2020) 022307.
[8] D. Helbing, I. Farkas, T. Vicsek, Nature 407 (2000) 487.
[9] D. Helbing, P. Molnar, I. J.Farkas, K. Bolay, Environ. Plan. B: Plann. Des. 28 (2001) 361.
[10] S.D. Dobrev, Soc. Sci. Res. 34 (2005) 800.
[11] N.A. Christakis, J.H. Fowler, Stat. Med. 32 (2013) 556.
[12] M. Belz, L.W. Pyritz, M. Boos, Behav. Process. 92 (2013) 6.
[13] P. Carpena, P. Bernaola-Galvan, P. Ch. Ivanov, Phys. Rev. Lett. 93 (2004) 176804.
[14] A.M. García-García, E. Cuevas, Phys. Rev. B 82 (2010) 033412.
[15] P. Carpena, P.A. Bernaola-Galván, C. Carretero-Campos, A.V. Coronado, Phys. Rev. E 94 (2016) 052302.