Reconstructing social sensitivity from evolution of content volume in Twitter

Sebastián Pinto,[*] Marcos A. Trevisan, and Pablo Balenzuela
Departamento de Física, FCEN, Universidad de Buenos Aires, Pabellón 1, Ciudad Universitaria, 1428EGA, Buenos Aires, Argentina and Instituto de Física de Buenos Aires, CONICET, Ciudad Universitaria, 1428EGA, Buenos Aires, Argentina

(Received 21 January 2022; accepted 4 November 2022; published 12 December 2022)

We set up a simple mathematical model for the dynamics of public interest in terms of media coverage and social interactions. We test the model on a series of events related to violence in the US during 2020, using the volume of tweets and retweets as a proxy of public interest, and the volume of news as a proxy of media coverage. The model successfully fits the data and allows inferring a measure of social sensibility that correlates with human mobility data. These findings suggest the basic ingredients and mechanisms that regulate social responses capable of igniting social mobilizations.

DOI: 10.1103/PhysRevE.106.064308

I. INTRODUCTION

The continuous expansion of the digital environment creates new and faster ways to exchange information and opinions [1]. At the same time, it also provides access to unprecedented amounts of data, allowing the quantitative investigation of the forces that underlie the diffusion of information [2] and the formation of public interest [3,4].

Dynamical systems have been particularly successful in identifying collective mechanisms that give rise to public opinion [5,6]. Using variables that describe the expansions and contractions of content volume, these models explain empirical data remarkably well [7].

In the domain of social media, the emergence of extreme opinions that arise from moderate initial conditions has been recently disclosed [8,9]. But extreme social reactions appear also beyond the domain of opinions and debates. Normally, people react to the news by sharing information and discussing opinions. In a few occasions, however, and under heightened social sensitivity, a reactive state may emerge giving rise to street manifestations, protests, and riots [10] that have been extensively studied and modeled [11,12].

Is it possible to extract a measure of social sensitivity from content volume coming from digital media? Here we hypothesize that the social sensitivity regulates the dynamics of the public reacting to the media coverage of massive events. To test this hypothesis, we set up a deliberately simple model for public interest modulated by media coverage and social interactions [13–16] that allows us to infer a measure of social sensitivity. We capitalize on the paradigmatic model developed by Granovetter [17] based on the concept of critical mass, which represents the fraction of interested people needed to induce interest to the rest of the population. We investigate this in connection with a series of highly sensitive events that took place in the US during 2020.

[*]spinto@df.uba.ar

II. DATA

The Black Lives Matter movement [18] encompasses events of different nature and volume of activity in the social media [Fig. 1(a)]. Here we analyze a subset of the events well covered by media sources, as displayed in chronological order in Fig. 1(b). The time evolution of these events is shown in Fig. 1(c). Representing the public interest, we show in black the volume of tweets and retweets containing the keywords George Floyd, Breonna Taylor, Jacob Blake, Rayshard Brooks, Ahmaud Arbery, and Andrés Guardado. Red filled curves correspond to the volume of tweets from the 29 most followed official media accounts containing the same keywords [19].

Besides a general resemblance of the public interest (black) and media coverage (red) across events, the traces are not merely copies of each other. One common feature is that the public interest grows faster than coverage at the events' onset. Here we propose that this effect is explained by the heightened social sensibility that characterizes these type of events. For this purpose we set up a model based on the one proposed by Granovetter, detailed in the next section.

III. MODEL

Our approach is grounded in the Granovetter model [17], originally proposed to explain the emergence of riots. In this model, agents adopt a binary state s, which we interpret as interested ($s = 1$) or noninterested ($s = 0$) in the event. The dynamics of the system is described in terms of the public interest, the fraction $p = \sum_i^N s_i/N$, where N is the size of the system. Each agent is characterized by a threshold τ_i, which is the fraction of interested agents needed to induce interest on the agent. Thresholds are random variables whose cumulative distribution $S(p) = P(\tau < p)$ is interpreted here as social engagement, given that it represents the fraction of agents that become active due to their threshold lies below p.

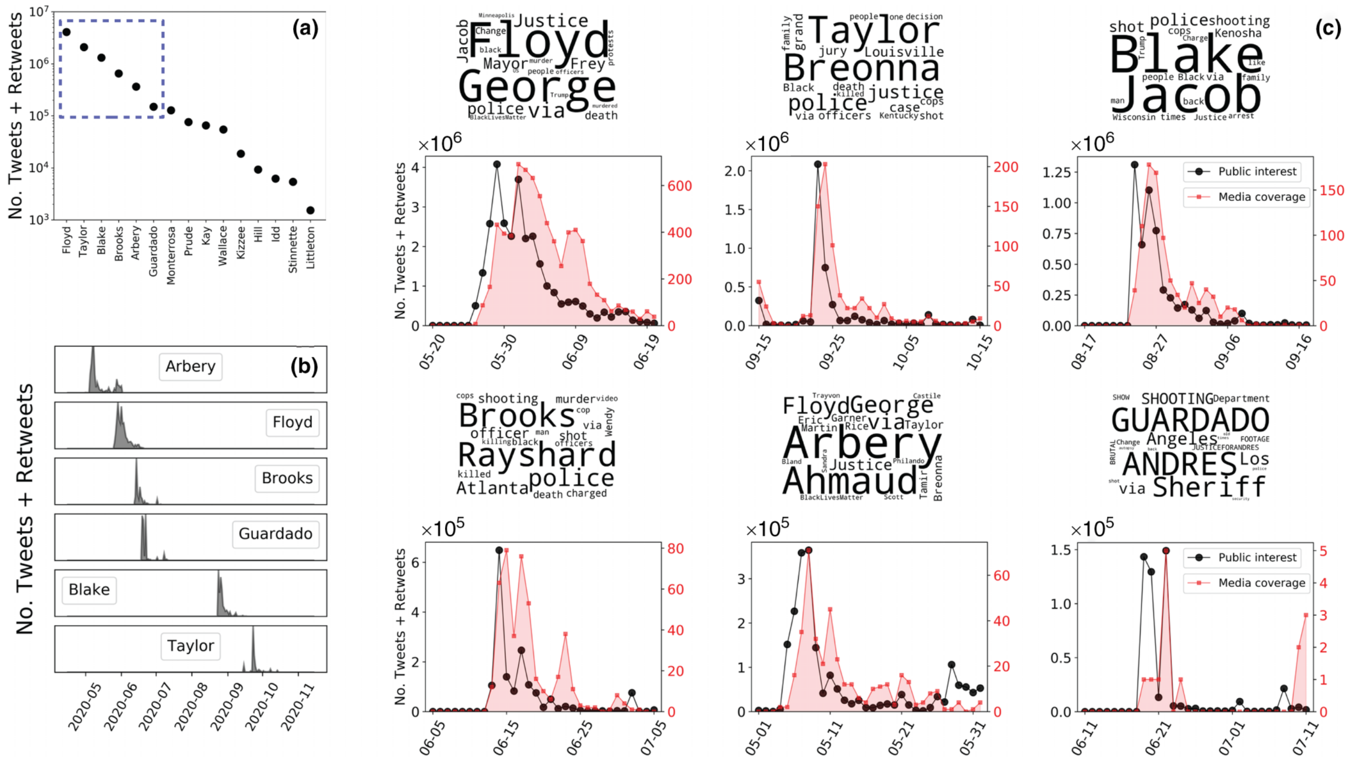

FIG. 1. Evolution of public interest and media coverage. (a) Peak values of tweets and retweets for events related to the Black Lives Matter movement. Dashed square shows those with enough statistics, limited by the amount of media tweets (in the case of Guardado is about ten tweets for the media accounts sampled). (b) Time traces of the volume of tweets and retweets, in chronological order. (c) Time traces of the volume of tweets and retweets [black circles, same data as (b)] and media accounts tweets (filled area).

Assuming that thresholds are normally distributed $\tau \sim N(\mu, \sigma)$, we have:

$$S(p|\mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{p} e^{-\frac{(\tau-\mu)^2}{2\sigma^2}} d\tau. \qquad (1)$$

When μ is low, small groups can trigger the interest to the rest of the system. On the contrary, high values of μ would require a bigger fraction of interested people to induce interest to the rest of the population. We therefore identify the quantity $1 - \mu$ as the social sensitivity of the population.

In his original model, Granovetter described the dynamics of the public interest p regardless of the influence of the media. To include this, we propose a modified model that reads (see details in Materials and Methods):

$$\frac{1}{\gamma}\frac{dp}{dt} = -p + e\,C(t) + (1-e)\,S[p|\mu(t), \sigma]. \qquad (2)$$

When the system is not exposed to the media ($e = 0$), we recover the original Granovetter model, in which the dynamics of the public interest p is driven by the social engagement S with a time scale controlled by γ. On the contrary, when exposure to the media is maximum ($e = 1$), the public interest is only driven by the media coverage C. In the general case $e \in (0, 1)$, media coverage acts as an external field that modulates the public interest. Of course, media coverage and public interest are far from being independent of each other. On the contrary, they feed one another; in mathematical terms, a closed model would require another equation for the evolution of media coverage modulated by the public interest. Here we tackle this by feeding Eq. (2) with the experimental time traces of media coverage $C(t)$.

Let us summarize the principal components of our model. On the one hand, we have two variables that quantify the volume of opinions and information shared by people: the public interest p and the media coverage C. On the other hand, we have the social sensitivity $(1 - \mu)$ and the social engagement S, two variables that describe macroscopic interactions among people. In the next section we show that the social variables can be reliably derived from the collected data shown in Fig. 1(c).

IV. RESULTS

To reconstruct the social variables, we integrate Eq. (2) using the volume of tweeted news as a proxy for the coverage $C(t)$. We seek for the functions $S(t)$ that minimize the difference between the resulting public interest and the volume of tweets and retweets. In Fig. 2 we show the best fitting curves for the public interest (top panels) and the reconstructed social engagement and social sensitivity (bottom panels, gray and red curves, respectively).

The two social variables are of a different nature. In fact, while the engagement $S(t)$ is a threshold-based variable whose dynamics can be expected to be fast, $1 - \mu(t)$ represents the slower, more gradual buildup of social sensitivity across the whole population. Accordingly, we find that this variable changes appreciably over periods of $\sim$ 15 days,

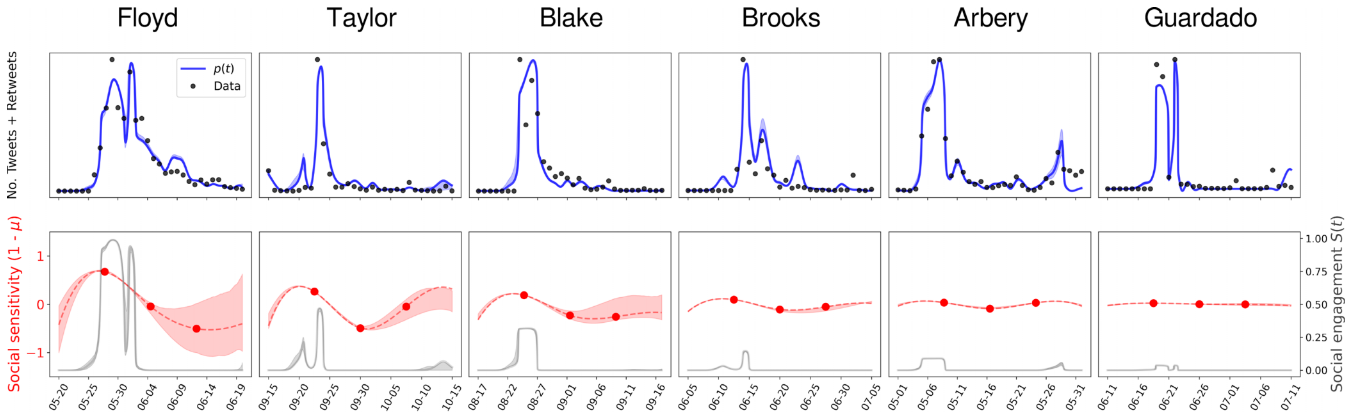

FIG. 2. Data fitting allows inferring social interactions. Top: Points correspond to public interest (tweets and retweets) along with the best fitting curves $p(t)$ (blue) obtained with the model of Eqs. (2) and (1). Bottom: Social sensitivity $1 - \mu(t)$, normalized to the event of mayor interest, which in this case is the murder of George Floyd. When the social sensitivity is high, more people become susceptible to the event. In gray lines the normalized social engagement $S(t) = S[p|\mu(t), \sigma]$ are also plotted.

which is, as expected, longer than the typical time scales of the media coverage and public interest (see Methods).

A summary of the fitting parameters is found in Table I. We find that exposure is rather stable across events, $e \sim 0.4$. This says that although media coverage is important, people are mainly influenced by the social environment in these kinds of events. Different from exposure, the time scale γ^{-1} decreases when the events accumulate over time. This is also expected, since the first four events (Arbery, Floyd, Brooks, and Guardado) occurred one immediately after the other [Fig. 1(b)], speeding up the dynamics of public interest along the sequence. After a pause of about two months, the same speeding up effect is seen for Taylor, which occurred right after Blake.

To quantify the performance of our model, we compare its goodness of fit with two basic models: one in which coverage is predicted by public interest alone, and the opposite one where public interest is predicted by coverage alone (see Methods). In Fig. 3 we show the mean-square errors for the three models. Comparison of the basic models shows that public interest tends to predict coverage better than coverage predicts public interest. This is also apparent from the time series [Fig. 1(c)], where the response of the media is slower with respect to the public interest at the onset of the events. Our model performs better than the basic models, explaining this delay by an increase in the social sensitivity $1 - \mu(t)$.

TABLE I. Fitted parameters. Events are in chronological order (Fig. 1). In all cases $\sigma = 0.2$. Intervals correspond to the 95% confidence levels (see Data Fitting section for more details).

($\times 10^{-3}$ days)	Exposure e	Timescale γ^{-1}
Arbery	0.35 (0.31, 0.39)	59 (36, 100)
Floyd	0.42 (0.40, 0.45)	36 (22, 59)
Brooks	0.47 (0.47, 0.56)	13 (12, 13)
Guardado	0.25 (0.25, 0.26)	10 (10, 16)
Blake	0.30 (0.26, 0.31)	22 (17, 100)
Taylor	0.47 (0.46, 0.54)	13 (10, 16)

The inferred dynamics of the social variables are shown in the bottom panels of Fig. 2.

Bottom panels of Fig. 2 show periods of time of increasing social sensitivity, which leads to a sudden increase of the social engagement, when a macroscopic fraction of agents becomes interested in the events. If this dynamics is accurate, we should expect an impact beyond the digital environment. To investigate the emergence of measurable collective activity associated to an increase in social sensitivity, we collected mobility measures across the US territory [20]. In Fig. 4 we show attendance to recreation places, groceries, pharmacies, and public transport stations in the counties and periods of time when the events took place. We find different degrees of correlation between the social sensitivity and mobility patterns for the most populous events using a lag of three days. In the case of Floyd, social sensitivity correlates with all the four mobility measures, with a peak in the mean Spearman's rank coefficient $r = -0.82$; in the case of Taylor, $r = -0.47$ for two of the mobility measures; for Blake, $r = -0.48$ and only one measure ($p < 0.05$ in all cases). The last three events were less massive, and we find no significant correlations with social sensitivity accordingly.

Taken together, these results suggest that our low-dimensional approximation of the Granovetter model captures the basic ingredients that regulate social responses of very different magnitudes, which are indeed capable of igniting social mobilizations. The model implements the hypothesis that agents become involved from media exposure and also from the presence of a critical mass of interested agents in the

TABLE II. Variables and parameters of the model.

Variable/Parameter	Interpretation
$p(t)$	Public interest
$C(t)$	Media coverage
$S(p(t))$	Social engagement
$1 - \mu(t)$	Social sensitivity
e	Media exposure
γ	Timescale

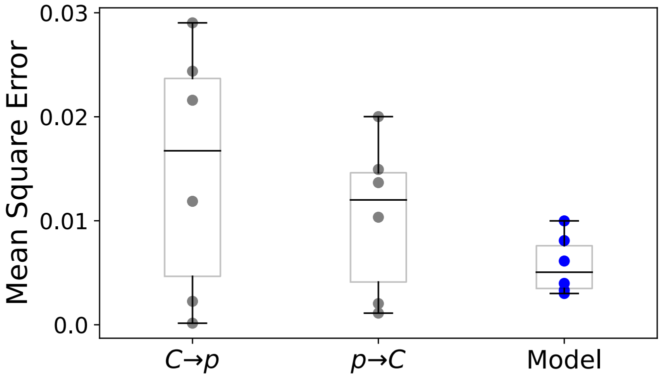

FIG. 3. Performance of the model. We compare the goodness of fit with two basic models across the six events analyzed here. In one model, public interest alone predicts coverage ($p \to C$) and in the other, coverage alone predicts public interest ($C \to p$). Our model explains the data better than both basic models with the same number of fitting parameters.

system, which leads to characterizing the social sensitivity of the population.

V. DISCUSSION AND CONCLUSIONS

Fluctuating interactions among people in massive social events are difficult to quantify. In this work we set up a simple mathematical model that allows us to infer how social interaction influences volume content representing public interest knowing media coverage. We then test our model on Twitter volume data related to the Black Lives Matter movement.

We find that this formulation fits the experimental series better than two models in which public interest and coverage explain each other, in absence of social interactions. Crucially, we show that the evolution of the social sensibility correlates with variations in mobility data due to protests and riots during the events that draw the majority of the attention, presumably the most moving ones.

A possible limitation of our model is related to the assumption of uniform mixing [21] in pairwise interaction, given that public interest time series were collected from Twitter, which is indeed highly structured. The topology of social networks plays a key role when dealing with opinions of different sign, which give rise, for instance, to the emergence of echo chambers [22,23]. In our work, however, we are dealing with the volume of keywords, regardless of ideological leanings. We show that, at least for the highly sensitive events analyzed here, the structure of the network can be disregarded, in line with similar models that assume uniform mixing and successfully explain the dynamics of time series related to different hashtags in Twitter [5–7]. Simple as it is, our model provides direct and interpretable measures of social engagement.

We are witnessing a rapid development of algorithms that are capable of organizing massive amounts of data based on statistical relationships. However, this growth has not been matched with a development of dynamical models capable of generalize our knowledge [24]. We hope that this work contributes to our understanding of public interest, showing the potential of a simple model to explain social reactions within and outside the digital environment.

VI. MATERIALS AND METHODS

A. Corpus of data

We collected all the available tweets in English containing the keywords George Floyd, Breonna Taylor, Jacob Blake, Rayshard Brooks, Ahmaud Arbery, Andrés Guardado, Sean Monterrosa, Daniel Prude, Deon Kay, Walter Wallace Jr., Dijon Kizzee, Andre Hill, Dolal Idd, Marcellis Stinnette, and Hakim Littleton, in a period of one month around a significant event related to each topic.

Tweets were collected using the Twitter API v2 [25]. We also collected the tweets with the same keywords from the group of most followed news accounts in Twitter [26]: @cnnbrk, @nytimes, @CNN, @BBCBreaking, @BBCWorld, @TheEconomist, @Reuters, @WSJ, @TIME, @ABC, @washingtonpost, @AP, @XHNews, @ndtv, @HuffPost, @BreakingNews, @guardian, @FinancialTimes, @SkyNews, @AJEnglish, @SkyNewsBreak, @Newsweek, @CNBC, @France24_en, @guardiannews, @RT_com, @Independent, @CBCNews, @Telegraph [27].

Mobility measures correspond to the US County associated to each event. From all mobility-related time series we extract the trend to compare with social engagement.

We provide here a brief context of the analyzed events. George Perry Floyd Jr. was murdered by a police officer in Minneapolis (Ramsey County), Minnesota, on May 25, 2020. Breonna Taylor was fatally shot in Louisville (Jefferson County), Kentucky, on March 13, 2020. On September 23, several protests occurred after charging decision announced in Taylor's death. Jacob S. Blake was shot and seriously injured by a police officer in Kenosha County, Wisconsin, on August 23. Rayshard Brooks was murdered on June 12, 2020 in Atlanta (Fulton County), Georgia. Ahmaud Arbery was murdered on February 23, 2020 in Glynn County, Georgia. The case became resonant after the viralization of a video about the shooting that caused his death on May 7. Andrés Guardado was killed by a deputy sheriff in Los Angeles County, California, on June 18, 2020.

B. Data fitting

We first normalized both public interest p and media coverage C respect to their peak values. To find a timescale for the dynamics of the social sensitivity, we parameterized $\mu(t)$ as a cubic spline of N equally spaced nodes within a one-month period. The fitting error either falls abruptly at $N = 5$ (Floyd and Blake) or does not change significantly in the range $4 \leqslant N \leqslant 9$ (Taylor, Brooks, Arbery, and Guardado). We therefore fixed the value $N = 5$, for which μ changes appreciably on a timescale of ∼15 days.

The media coverage was interpolated in order to obtain a continuous signal. Interpolation and numerical integration of Eqs. (1) and (2) were performed with the library SCIPY [28]. Parameter fitting was performed using a grid search in parameter $\gamma \in [10^{-1}–120]$ in combination with a minimization routine for a the rest of the parameters ($e \in [0, 1]$ and nodes of $\mu \in [-1, 2]$). The routine consists in integrating the

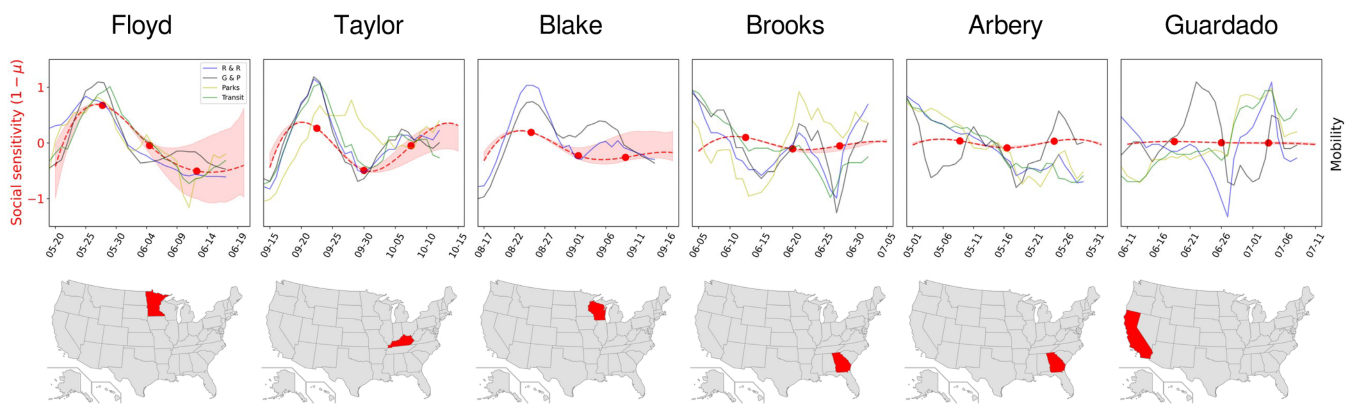

FIG. 4. Correlations between social sensitivity and mobility patterns. Social sensitivity (red) and standardized mobility observables of the corresponding county. R & R: retail and recreation; G & P: groceries and pharmacies; Parks: public parks; Transit: transit in public transport stations (Parks and Transit not shown for Blake due to lack of data). All mobility measures were shifted −3 days and inverted for visualization purposes.

model and varying the parameters until a convergence criteria is reached. We used sequential least-squares programming for bounded problems in SCIPY to minimize the mean-square error between the output of the model and data.

Confidence intervals provided in Table I and shown in Figs. 2 and 4 correspond to fitting solutions with an error up to 10% of the best solution in each case, except for Taylor and Brooks, where solutions with a fitting error up to 50% of the best solution were reported.

C. Basic models

We compare the goodness of fit with two basic models. In one of them, coverage is predicted by public interest $p \rightarrow C$ and in the other it is the other way around, $C \rightarrow p$. Both basic models were set up to be nonlinear functions approximated by order seventh polynomials, $C(t) = \sum_{n=1}^{7} a_n\, p^n(t)$ and $p(t) = \sum_{n=1}^{7} b_n\, C^n(t)$, without zeroth-order term ($a_0 = b_0 = 0$). In this way, the basic models match the number of fitting parameters of the model [e, γ and $\mu(t_i)$, with $1 < i < 5$].

D. Analytical formulation of the model

Equation (1) is an analytical approximation of the threshold-based model proposed by Granovetter [17] with the addition of an external field. In this model, agents adopt a binary state s, which we interpret as interest ($s = 1$) or noninterest ($s = 0$) in a given topic. The dynamics of the system is described in terms of the fraction of interested agents $p = \sum_i s_i/N$, where N is the size of the system. The agents have also an associated threshold τ_i, which is the fraction of interested agents needed to induce interest on agent i. The thresholds are random variables between 0 and 1 taken from a probability density $f(\tau)$. On the other hand, the external field is introduced through a parameter $C \in [0, 1]$ independent of the state of the system.

With these ingredients, the dynamics of the system is as follows: the fraction of interested agents p can change because a random agent i interacts with the media with probability e and become interested ($s_i = 1$) in a given topic with probability C or disinterested ($s_i = 0$) with probability $1 - C$; otherwise, with probability $1 - e$, the agent observes the system. In this last case, if the fraction of interested agents is greater than the threshold of the agent ($p \geqslant \tau_i$), then it becomes interested ($s_i = 1$); otherwise, it becomes disinterested ($s_i = 0$). Agents' states are synchronously updated, independently from their initial state.

Following Ref. [29], we derive the analytical expression for the dynamics of p shown in Eq. (1). Let $q(p_k, t)$ be the probability that the fraction of interested agents at time t is $p_k = k/N$. The master equation for $q(p_k, t)$ is

$$\frac{dq(p_k,t)}{dt} = Q(1|p_{k-1})q(p_{k-1},t) + Q(0|p_{k+1})q(p_{k+1},t) - Q(1|p_k)q(p_k,t) - Q(0|p_k)q(p_k,t),$$

where $Q(1|p_k)$ y $Q(0|p_k)$ are the transition probabilities that a given agent become interested or disinterested given p_k. These probabilities are given by:

$$Q(1|p_k) = (1 - p_k)[(1 - e)S(p_k) + eC]$$
$$Q(0|p_k) = p_k[(1 - e)(1 - S(p_k)) + e(1 - C)],$$

where $S(p_k)$ is the threshold cumulative distribution function $S(p_k) = \int^{p_k} f(\tau)d\tau$, which by definition is the fraction of agents whose threshold is below p_k $[S(p_k) \equiv P(\tau < p_k)]$.

In the limit of infinite population ($N \rightarrow \infty$), $p_k \rightarrow p$, where p is now the fraction of interested agents and a continue variable $\in [0, 1]$. In this limit, the following approximations are taken:

$$p_{k\pm1} \rightarrow p \pm \eta$$
$$q(p_{k\pm1}, t) \rightarrow q(p,t) \pm \frac{\partial q(p,t)}{\partial p}\eta$$
$$S(p_{k\pm1}) \rightarrow S(p) \pm \frac{\partial S(p)}{\partial p}\eta$$

with $\eta = 1/N$. Replacing the above expressions in the master equation and neglecting terms of η^2 order, we obtain:

$$\frac{\partial q(p)}{\partial t} = -\frac{\partial}{\partial p}[(-p + S(p) - eS(p) + eC)q(p,t)]\eta.$$

For a well-defined initial condition, $q(p,0) = \delta(p - p_0)$ [$\delta(x)$ is the Dirac's delta] and rescaling time $t \rightarrow Nt$, the solution of the above equation (pages 53–54 of Ref. [30]) is given by:

$$\frac{dp}{dt} = -p + (1-e)S(p) + eC.$$

In particular, if the thresholds are normally distributed with mean μ and dispersion σ, $S(p) \equiv S(p|\mu,\sigma)$. Finally, by adding a constant γ that allows to adjust the timescale, Eq. (1) is obtained.

Equation (1) has equilibria given by $p_{eq} = (1 - e)S(p_{eq}) + eC$ The stability of these points is given by the sign of:

$$\frac{d\dot{p}}{dp}|_{p_{eq}} = \left[-1 + (1-e)\frac{dS(p)}{dp}\right]_{p_{eq}},$$

where it can be observed that the parameter C plays no role in setting the stability. As reference, we summarize here all the variables and parameters of the model mentioned during the paper:

ACKNOWLEDGMENTS

This research was partially funded by the Universidad de Buenos Aires (UBA), the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) through Grant No. PIP-11220200102083CO, and the Agencia Nacional de Promoción de la Investigación, el Desarrollo Tecnológico y la Innovación through Grant No. PICT-2020-SERIEA-00966.

[1] F. Wu and B. A. Huberman, Novelty and collective attention, Proc. Natl. Acad. Sci. **104**, 17599, (2007).

[2] J. Leskovec, L. Backstrom, and J. Kleinberg, Meme-tracking and the dynamics of the news cycle, in *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (ACM, 2009), pp. 497–506.

[3] S. Pinto, F. Albanese, C. O. Dorso, and P. Balenzuela, Quantifying time-dependent media agenda and public opinion by topic modeling, Physica A **524**, 614 (2019).

[4] F. Albanese, S. Pinto, V. Semeshenko, and P. Balenzuela, Analyzing mass media influence using natural language processing and time series analysis, J. Phys.: Complex. **1**, 025005 (2020).

[5] S. Towers, S. Afzal, G. Bernal, N. Bliss, S. Brown, B. Espinoza, J. Jackson, J. Judson-Garcia, M. Khan, M. Lin, *et al.*, Mass media and the contagion of fear: The case of ebola in america, PLoS ONE **10**, e0129179 (2015).

[6] M. Muhlmeyer, J. Huang, and S. Agarwal, Event triggered social media chatter: A new modeling framework, IEEE Trans. Comput. Soc. Syst. **6**, 197 (2019).

[7] P. Lorenz-Spreen, B. M. Mønsted, P. Hövel, and S. Lehmann, Accelerating dynamics of collective attention, Nature Commun. **10**, 1759 (2019).

[8] F. Baumann, P. Lorenz-Spreen, I. M. Sokolov, and M. Starnini, Phys. Rev. X **11**, 011012 (2021).

[9] F. Baumann, P. Lorenz-Spreen, I. M. Sokolov, and M. Starnini, Modeling Echo Chambers and Polarization Dynamics in Social Networks, Phys. Rev. Lett. **124**, 048301 (2020).

[10] J. Drury, C. Stott, R. Ball, S. Reicher, F. Neville, L. Bell, M. Biddlestone, S. Choudhury, M. Lovell, and C. Ryan, A social identity model of riot diffusion: From injustice to empowerment in the 2011 London riots, Eur. J. Soc. Psychol. **50**, 646 (2020).

[11] M. Agamennone, Riots and uprisings modelling Conflict between centralised and decentralised systems, Ph.D. thesis, King's College London, 2020.

[12] L. Bonnasse-Gahot, H. Berestycki, M. A. Depuiset, M. B. Gordon, S. Roché, N. Rodriguez, and J. P. Nadal, Epidemiological modelling of the 2005 French riots: A spreading wave and the role of contagion, Sci. Rep. **8**, 107 (2018).

[13] L. Guo and M. McCombs, *The Power of Information Networks: New Directions for Agenda Setting* (Routledge, Oxfordshire, 2015).

[14] C. Castellano, S. Fortunato, and V. Loreto, Statistical physics of social dynamics, Rev. Mod. Phys. **81**, 591 (2009).

[15] P. Balenzuela, J. P. Pinasco, and V. Semeshenko, The undecided have the key: Interaction-driven opinion dynamics in a three state model, PLoS ONE **10**, e0139572 (2015).

[16] F. Barrera Lemarchand, V. Semeshenko, J. Navajas, and P. Balenzuela, Polarizing crowds: Consensus and bipolarization in a persuasive arguments model, Chaos **30**, 063141 (2020).

[17] M. Granovetter, Threshold models of collective behavior, Am. J. Soc. **83**, 1420 (1978).

[18] N. Carney, All Lives Matter, but so Does Race, Humanity Soc. **40**, 180 (2016).

[19] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.106.064308 for further details in the quantifying of media coverage.

[20] https://www.google.com/covid19/mobility/.

[21] K. Burghardt, W. Rand, and M. Girvan, Inferring models of opinion dynamics from aggregated jury data, PLOS ONE **14**, e0218312 (2019).

[22] M. Cinelli, G. De Francisci Morales, A. Galeazzi, W. Quattrociocchi, and M. Starnini, The echo chamber effect on social media, Proc. Natl. Acad. Sci. **118**, e2023301118 (2021).

[23] W. Cota, S. Ferreira, R. Pastor-Satorras, and M. Starnini, Quantifying echo chamber effects in information spreading over political communication networks, EPJ Data Sci. **8**, 35 (2019).

[24] S. L. Brunton, J. L. Proctor, J. N. Kutz, and W. Bialek, Discovering governing equations from data by sparse identification of nonlinear dynamical systems, Proc. Natl. Acad. Sci. USA **113**, 3932 (2016).

[25] https://developer.twitter.com/en/use-cases/do-research/academic-research.

[26] https://www.intelligencefusion.co.uk/insights/resources/article/top-30-most-followed-news-accounts-on-twitter/.

[27] Twitter data is available at https://shorturl.ae/AcUge.

[28] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey *et al.*, SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python, Nature Methods **17**, 261 (2020).

[29] A. R. Akhmetzhanov, L. Worden, and J. Dushoff, Effects of mixing in threshold models of social behavior, Phys. Rev. E **88**, 012816 (2013).

[30] C. W. Gardiner *et al.*, *Handbook of Stochastic Methods*, Vol. 3 (Springer, Berlin, 1985).