Universality and correlations in individuals wandering through an online extremist space

Z. Cao, M. Zheng, Y. Vorobyeva, C. Song, and N. F. Johnson

Phys. Rev. E **97**, 032315 — Published 26 March 2018

DOI: 10.1103/PhysRevE.97.032315

Universality and correlations in individuals wandering through an online extremist space

Z. Cao[1], M. Zheng[1], Y. Vorobyeva[2], C. Song[1], N.F. Johnson[1]
[1] *Physics Department, University of Miami, Coral Gables, FL 33146, U.S.A. and*
[2] *Dept. of International Studies, University of Miami, Coral Gables, FL 33146, U.S.A.*
(Dated: March 13, 2018)

The 'out of the blue' nature of recent terror attacks and the diversity of apparent motives, highlight the importance of understanding the online trajectories that individuals follow prior to developing high levels of extremist support. Here we show that the physics of stochastic walks, with and without temporal correlation, provides a unifying description of these online trajectories. Our unique dataset comprising all users of a global social media site, reveals universal characteristics in individuals' online lifetimes. Our accompanying theory generates analytical and numerical solutions that describe the characteristics shown by individuals that go on to develop high levels of extremist support, and those that do not. Going forward, it is conceivable that a deeper understanding of these temporal and many-body correlations may eventually contribute to the important task of better countering the spread of radical propaganda online.

I. INTRODUCTION

Following the terror attacks in London, Manchester, Washington D.C. and Paris in 2017, and Orlando, Berlin, Nice and Brussels in 2016, authorities face the fundamental problem of detecting individuals who are currently developing intent in the form of strong support for some extremist entity – even if they never end up doing anything in the real world. The importance of online connectivity in developing intent [1–8] has been confirmed by case-studies of already convicted terrorists by Gill and others [1–5]. Quantifying this online dynamical development can help move beyond static watch-list identifiers such as ethnic background or immigration status. Heuristically, one might imagine that an individual who enters an online space, wanders through the content available and – depending in part on what they find on any given day – feels pulled toward, or pushed away from, a particular extreme ideology. This process of individual fluctuation will be made even more complex by endogenous and exogenous factors in their own lives. Adding to the complication, humans are heterogeneous and hence may enter an online space at different times, spend different amounts of time online, and may end up losing interest and dropping out entirely, continuing in an uncertain state, or developing a high level of support.

We show here that despite this wide range of possible behaviors, a surprising level of universality arises in the online trajectories of individuals through an extremist space. We provide a stochastic walk model that connects together all individuals, even though they may end up with very different outcomes. Though our focus is on individuals' online dynamics irrespective of whether they later carry out an extremist act or not, subsequent analysis of media reports together with others' postings suggest that a significant number of individuals in our dataset do. Our dataset is assembled using the same methodology as Ref. [9], and includes the global population of ∼ 350 million users of the social media outlet VKontakte (www.vk.com) which became the primary online social media source for ISIS propaganda and recruiting during 2015 [9]. Unlike on Facebook where pro-ISIS activity is almost immediately blocked, support on VKontakte develops around online groups (i.e. self-organized communities) which are akin to Facebook groups that support everyday topics such as a sport team.

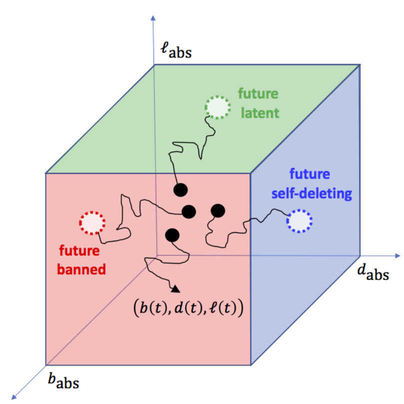

FIG. 1. (Color online) Schematic of possible individual trajectories in $d-\ell-b$ space bounded by absorbing barriers at $d_{\rm abs}, \ell_{\rm abs}, b_{\rm abs}$ such that $d(t) < d_{\rm abs}$, $\ell(t) < \ell_{\rm abs}$ and $b(t) < b_{\rm abs}$. These illustrate the three possible outcomes of interest.

II. THEORY

A. Clock-time Lifetime

Given that other forms of extremism ranging from far-left to far-right *also* appear through such online groups (e.g. the Washington D.C. shooter [10] and Maryland attacker [11] were both members of such groups on Facebook), and given that many social media sites allow such community features (i.e. online groups), our results and model should have general applicability. Even the encrypted application Telegram allows users to set up 'super-groups' [8]. All these online groups tend to keep themselves open-source in order to attract new members, hence we were able to record the current membership of pro-ISIS groups at every instant using entirely open-source information. Each individual moving through such an online space can be classified at any time t by what will happen to them *in the future*, even though he/she may at time t still be undecided about supporting the ideology, or may even be moving away from it. Each individual at time t has one of four unique labels:

Future banned: At some future time, he/she will develop and hence express such a high level of extremist support that their account will get banned by moderators. These individuals would likely be of *most* interest to authorities.

Future latent: At some future time, he/she will stop being a member of extremist groups but will not self-delete their account, perhaps reflecting indifference to the extremist ideology.

Future self-deleting: At some future time, he/she self-deletes their own account, perhaps because they are scared of being tracked.

Still ongoing: He/she will remain in development. Their account remains unbanned and they continue joining pro-ISIS groups (and possibly leaving, though group leaving events are rare).

We focus here on the first three individual types since they provide us with well-defined lifetimes and timelines in terms of the online extremist groups that they join. Banning and self-deleting events are announced on a user's webpage by moderators when they occur. The clock-time lifetimes are $T_{\rm ban}$, $T_{\rm lat}$, and $T_{\rm s-del}$: $T_{\rm ban}$ is the time interval between them first joining a pro-ISIS group and their account being banned; $T_{\rm lat}$ is the time interval between them first joining a pro-ISIS group and them ceasing to be a member of any pro-ISIS group; and $T_{\rm s-del}$ is the time interval between them first joining a pro-ISIS group and them self-deleting their account. The event-time lifetimes $L_{\rm ban}$, $L_{\rm lat}$, and $L_{\rm s-del}$, are given by the total number of groups that the individual joins during the observation period. The event-time lifetime for each individual is therefore an integer, and could if desired be converted to a physical time by multiplying by the average number of days it takes the individual to join one new group. We model the instant of banning as an individual hitting an absorbing barrier for the first time at $b_{\rm abs}$ in a one-dimensional walk $b(t)$, where $b(t)$ represents the level of extremism (i.e. pro-ISIS support) that an individual expresses. The instant of becoming latent is when an individual hits an absorbing barrier $\ell_{\rm abs}$ during a one-dimensional walk $\ell(t)$, where $\ell(t)$ represents the desire to become latent. The instant of self-deletion is when an individual hits an absorbing barrier at $d_{\rm abs}$ in a one-dimensional walk $d(t)$, where $d(t)$ represents an individual's desire to self-delete. Though an obvious over-simplification, such a single scalar parameter has already been adopted in other sociological contexts to mimic aspects of human personality [12, 13].

Each of the $\sim$ 350 million VKontakte users undergoes their own walk in the three-dimensional d-ℓ-b space in Fig. 1, characterized by the position vector $(b(t), \ell(t), d(t))$ and with absorbing barriers at $d_{\rm abs}, \ell_{\rm abs}, b_{\rm abs}$ such that $d(t) < d_{\rm abs}$, $\ell(t) < \ell_{\rm abs}$ and $b(t) < b_{\rm abs}$. We identify 7,707 individuals that eventually hit the barrier along the b-axis in Fig. 1 (i.e. future banned individuals); 65,169 that eventually hit the barrier along the ℓ-axis (i.e. future latent individuals); and 18,905 individuals that eventually hit the barrier along the d-axis (i.e. future self-deleting individuals).

Two questions then arise about what *type* of walk would be appropriate to describe each individual in Fig. 1, and how this might then be modeled in some approximate yet tractable way that will allow mathematical analysis for the entire population [14–21]. The correct answer to the first question would undoubtedly depend on the changing situations to which each individual is exposed in their daily life, their past experiences, and also their own unique personality – which is information to which we do not have access, therefore making the task of answering the second question a seemingly impossible one. However it has already been shown for other areas of human behavior [14, 15, 22–24] that when seeking an average description over an entire population, the walk that each individual takes does not have to be an accurate reflection of what they actually do at each timestep, but rather it needs to capture the stochastic nature over the whole population as each individual gets feedback from their own daily life and hence moves toward or away from a given barrier in their own particular way in Fig. 1. In this regard, we take inspiration from the somewhat similar situation of heterogeneous individuals operating in a financial market, where individuals decide separately each day whether to move to a position that favors an upward move in the market index (which could be considered analogous to moving toward the barrier in Fig. 1) or to one that favors a downward move (analogous to moving away from the barrier in Fig. 1). Despite the complex nature of the arrival of exogenous information, it is well-known that as a crude approximation the behavior of market participants can be captured by a *stochastic* walk which mimics the complex daily changes in their personal positions [15, 22–24]. As a result, even though the details of each individual's behavior are grossly over-simplified, the overall market index varies as if each individual were undergoing its own stochastic walk [22–24]. Indeed, the simplest description is that of a quasi-random walk [23, 24]. In a similar way, as a first attempt at describing online extremism, we will in this paper adopt a similar model in which individuals follow a simple stochastic walk. On any given day, therefore, a given individual either moves toward or away from the barriers in Fig. 1. Even with this gross simplification, the mathematical formulation of an individual's trajectory in Fig. 1 is still in principle daunting since, a priori, the components of the walks along each direction could be coupled for any given individual. However for simplicity, we will treat each individual as executing a $1+1+1$-dimensional walk [25]. As we will see in Fig. 2, the errors in so doing effectively wash out at the average population level, making this simple stochastic walk description a surprisingly accurate zeroth-order approximation – apart from the appearance of temporal correlations at short lifetimes (Fig. 2). In future work, we hope to tease out more details about these walks and relate them to the past content of the groups that individuals joined, or the content that they themselves posted, as well as the nature of other endogenous factors and exogenous stimuli to which they were exposed [16–21].

With this assumed decoupling in place, solving for the observational period T in a generic single dimension $x(t)$ with a single absorbing barrier at x_b solves the problem for each of these dimensions and yields a lifetime distribution for the entire process. Allowing for the possibility of a non-zero drift velocity u towards the respective barrier x_b, the Fokker-Planck equation for any component $x(t)$ in Fig. 1 becomes:

$$\begin{cases} (\frac{\partial}{dt} - D\frac{\partial^2}{\partial x^2} + u\frac{\partial}{\partial x})G(x,t;x_0,t_0) = \delta(x-x_0)\delta(t-t_0) \\ G(x_b,t;x_0,t_0) = 0 \end{cases} \tag{1}$$

where $x \leq x_b$, $0 \leq t - t_0 \leq T$. T is the observation period and D is the diffusion coefficient, assumed to be time-independent. For the simulations, we consider the discrete version with unit diffusion speed $\Delta x/\Delta t = 1$, with $D = (1-u^2)/2$ and $u = 2p - 1$ where p is the probability of moving forward at each timestep. The solution is $G(x,t;x_0,t_0) = \Theta(t-t_0)K(x,t;x_0,t_0)$, where the propagator $K(x,t;x_0,t_0) = \Phi(x-x') - \exp[-u(x_b-x)/D]\Phi[x-(2x_b-x')]$, $x' = x_0 + u(t-t_0)$, and $\Phi(x) = \exp[-x^2/(4Dt)]/\sqrt{4\pi Dt}$. To mimic human heterogeneity, we consider a uniformly distributed initial condition at $t = t_0 = 0$:

$$f_{1D}(x,t) = \begin{cases} \delta_{t,0}/x_m & x_b - x_m \leq x < x_b \\ 0 & \text{elsewhere} \end{cases} \tag{2}$$

where δ is the Kronecker delta and x_m is a normalization constant. $x_m = T\Delta x/\Delta t$ in the simulations. This is reasonable since individuals located below $x = x_b - x_m$ can never reach the boundary and hence can be ignored. The probability distribution

$$\begin{aligned} P_{1D}(x,t) &= \sum_{t_0=0}^{T}\int_{x_b-x_m}^{x_b} dx_0 K(x,t;x_0,t_0)f_{1D}(x_0,t_0) \\ &= \frac{1}{2x_m}\Bigg\{\psi\Big(\frac{x_b+ut-x}{\sqrt{4Dt}}\Big) + \psi\Big(\frac{x_m+x-x_b-ut}{\sqrt{4Dt}}\Big) \\ &- \exp\left[\frac{-u(x_b-x)}{D}\right]\Bigg[\psi\Big(\frac{x_b+x_m-ut-x}{\sqrt{4Dt}}\Big) \\ &+ \psi\Big(\frac{x+ut-x_b}{\sqrt{4Dt}}\Big)\Bigg]\Bigg\}, \end{aligned}$$

where $\psi(x)$ is the error function; and the total probability

$$\begin{aligned} R_S(t) &= \int_{-\infty}^{x_b} dx P_{1D}(x,t) \\ &= \frac{1}{2x_m u}\Bigg\{\left(D - x_m u + tu^2\right)\Psi\left(\frac{x_m - tu}{2\sqrt{Dt}}\right) \\ &- D\exp\left(\frac{x_m u}{D}\right)\Psi\left(\frac{x_m + tu}{2\sqrt{Dt}}\right) \\ &+ 2u\sqrt{\frac{Dt}{\pi}}\left[e^{-\frac{(x_m-tu)^2}{4Dt}} - e^{-\frac{tu^2}{4D}}\right] \\ &- \left(2D + tu^2\right)\psi\left(\frac{u}{2}\sqrt{\frac{t}{D}}\right) + 2x_m u - tu^2\Bigg\}, \end{aligned}$$

where $\Psi(x)$ is the complementary error function. The distribution of clock-time lifetimes

$$F(t) = \frac{-\frac{dR_S(t)}{dt}}{-\int_0^T dt\frac{dR_S(t)}{dt}} = \frac{-\frac{dR_S(t)}{dt}}{R_S(0) - R_S(T)}, \tag{3}$$

from which we obtain

$$\begin{aligned} F(t) = Z^{-1}\Bigg\{u\Bigg[\psi\Big(\frac{x_m - tu}{2\sqrt{Dt}}\Big) + \psi\left(\frac{u}{2}\sqrt{\frac{t}{D}}\right)\Bigg] \\ + \sqrt{\frac{4D}{\pi t}}\left[e^{-\frac{tu^2}{4D}} - e^{-\frac{(x_m-tu)^2}{4Dt}}\right]\Bigg\}, \end{aligned} \tag{4}$$

where the normalization constant

$$\begin{aligned} Z = \left(\frac{2D}{u} + Tu\right)\psi\left(\frac{1}{2}u\sqrt{\frac{T}{D}}\right) + \frac{D}{u}e^{\frac{x_m u}{D}}\Psi\left(\frac{x_m + Tu}{2\sqrt{DT}}\right) \\ + \left(x_m - Tu - \frac{D}{u}\right)\Psi\left(\frac{x_m - Tu}{2\sqrt{DT}}\right) + Tu \\ + 2\sqrt{\frac{DT}{\pi}}\left[e^{-\frac{Tu^2}{4D}} - e^{-\frac{(x_m-Tu)^2}{4DT}}\right]. \end{aligned}$$

When $u \to 0$,

$$F(t)|_{u\to 0} = Z_0^{-1}\sqrt{\frac{D}{\pi t x_m^2}}\left[1 - \exp\left(-\frac{x_m^2}{4Dt}\right)\right]; \tag{5}$$

and therefore

$$F(t)|_{u\to 0;\ t\ll x_m^2/(4D)} \sim t^{-1/2}, \tag{6}$$

where the normalization constant

$$Z_0 = 1 + \sqrt{\frac{4DT}{\pi x_m^2}}\left[1 - \exp\left(-\frac{x_m^2}{4DT}\right)\right] - \psi\left(\frac{x_m}{\sqrt{4DT}}\right).$$

Hence our theory predicts an approximate power-law distribution for clock-time lifetimes that are not too large, with a negative scaling exponent of magnitude $1/2$, which is much smaller than that of the first-hitting time of conventional 1-D Brownian motions ($\geq 3/2$, cf. [21]), indicating the existence of high heterogeneity in users' initial condition.

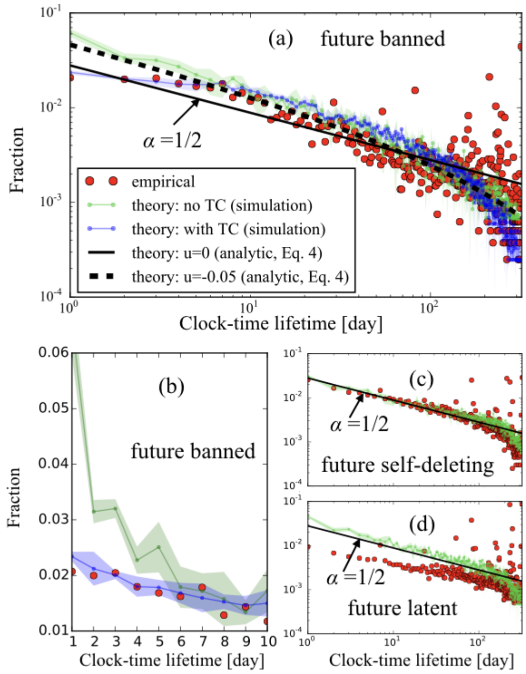

FIG. 2. (Color online) (a), (c) and (d) show the distribution of clock-time lifetimes for individuals with the three types of future outcomes shown in Fig. 1, together with our theoretical predictions. The legend in (a) applies to all. (b) enlarges (a) for short lifetimes. One standard deviation error bands are shown. The results with the TC (i.e. temporal correlation) effect for (c) and (d) are similar to the respective no-TC case, and hence are not shown.

Figures 2(a), (c) and (d) show that despite their very different origins and meanings, all three distributions tend to follow the same analytical 1/2 power-law for intermediate clock-time lifetimes. Moreover this agreement can be improved by adding a small u in Eq. (4) (e.g. Fig. 2(a)). A full numerical simulation of our model yields even better overall agreement (green curves). Deviations arise at short clock-time lifetimes for future-banned individuals (Figs. 2(a)-(b)). However the good agreement can be restored if we add temporal correlations (TC, i.e. memory) to our walk model: with probability q, an individual changes his/her $x(t)$ value at time t by adopting the same change that occurred in an earlier timestep randomly chosen from the last m timesteps. Even the simplest case of $m = 1$ shows good agreement (blue curves in Fig. 2).

B. Event-time Lifetime

Consistent with previous work suggesting that people are highly heterogeneous in how long they take to do something [16], we find that the empirical event-time lifetime can be quite different from the corresponding clock-time lifetime. This motivates us to look at event-time.

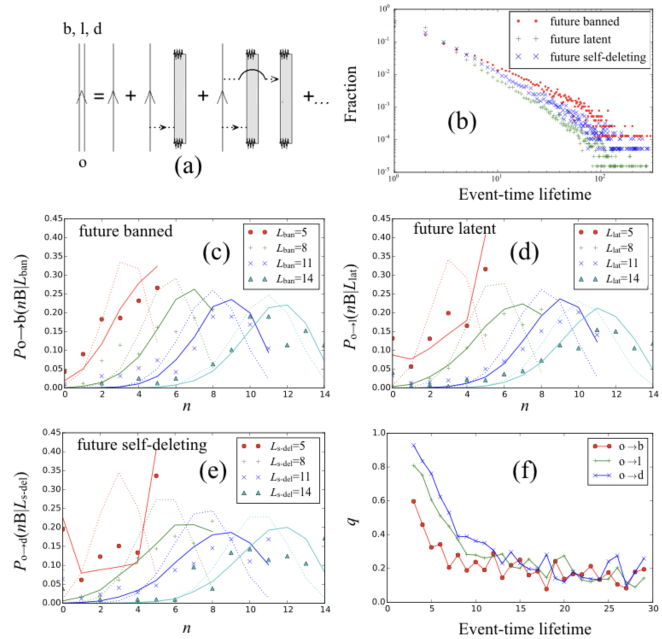

FIG. 3. (Color online) (a) Exact diagrammatic expansion [26] of the probability that a randomly chosen individual ends up as banned (denoted as 'b'), self-deleting ('d') or latent ('l') after joining $L = 0, 1, 2, \ldots$ online groups of any type. $L = 0, 1, 2$ terms are shown explicitly. (b) Terms from (a) evaluated empirically by counting the fraction of individuals who end up as 'b','d' or 'l' after joining exactly L online groups (i.e. event-time lifetime is L). (c)-(e) Expansion terms from (a) for a specific event-time lifetime (different subscripts for L are used to distinguish the user types), where here n only counts the number of future-banned groups (B) joined (see text) and hence the maximum n appears bounded from above by the event-time lifetime. Points are empirical results. Solid lines are our temporal correlation (TC, i.e. memory) model results determined using Maximum Likelihood Estimation (MLE). Dashed lines are null model (i.e. binomial distribution). (f) MLE q values vs. event-time lifetime in our model. MLE p value ≈ 0.73 in all cases.

We represent the probability of an individual ending up in one of the three possible outcomes in Fig. 1, as an expansion with respect to L (Fig. 3(a)) where each term is the probability that this (i.e., an individual ending up in one of the three possible outcomes) happens (i.e. lifetime ends) after joining L online groups [26] (L is the event-time lifetime). Figure 3(b) shows the distributions of these empirically determined expansion terms. The fact that the distribution for future banned users is the broadest, suggests that future banned users tend to be the most active in joining groups.

In order to check further if the distributions are power-law distributions, we tried fitting the three empirical distributions using a (discrete) power law distribution, following the standard maximum-likelihood method of Clauset et al. [27]. We took the minimum event-time lifetime for fitting to be 1. We found the estimated power-law exponents to be α =2.1, 1.6, and 1.9 for the future latent, future banned, and future self-deleting users, respectively. However, the p-value in each case is small (< 0.05) which suggests that a power-law form can be rejected with high confidence [27]. Hence, there is no strong statistical support for a power law.

Instead, to look for meaningful patterns, we move to conditional probabilities. In the following analysis, we attach different subscripts to L in order to distinguish the user types (i.e., $L_{\rm ban}$ for future banned users, $L_{\rm lat}$ for future latent users, and $L_{\rm s-del}$ for future self-deleting users. Figures 3(c)-(e) shows results conditioned on the event-time lifetime from (b) and counting as n the number of future-banned groups joined during the lifetime, i.e. n in Figs. (c)-(e) only counts groups who will themselves get banned by moderators. These are of the most interest since by definition they will develop the most extreme content.

Just as for clock-time, Figs. 3(c)-(e) show that a stochastic walk model without temporal correlations (TC) provides poor agreement for short event-time lifetimes. Similar to before, we therefore introduce TC and hence memory: At each step, with probability q the individual decides to join a group of the same type (i.e. either future-banned or not) as they did in one of their last m joining events, randomly chosen from m. Hence with probability $(1-q)p$, they join a future-banned group, and with probability $(1-q)(1-p)$ they join a non future-banned group. m acts as a memory length, q is the probability of making a decision according to this memory, and p determines individual preference for a specific group type. The model simulation is over 10,000 individuals. q is the dominant parameter in determining the model fit.

As an illustration, we describe here the analysis for the future-banned users. Stochastic simulations show that increasing m strengthens the memory effect significantly only when q is sufficiently large (e.g. above ~ 0.7); therefore, for most values of q, the profile of the distribution $P_{o\to b}(n\mathrm{B}|L_{\rm ban})$ is primarily determined by q. Hence we let $m = 1$ for simplicity, and estimate q and p for each value of $L_{\rm ban}$ from the empirical data using maximum likelihood estimation (MLE), and perform a model fit to the empirical results by simulation. For the memory model with $m = 1$, the likelihood for an individual i to have a path $\mathbb{S}_i = \{S_i[t]|S_i[t] \in \{0, 1\}, t = 0, 1, 2, ..., L-1\}$ ($S_i[t] = 0$ corresponds to joining a future-banned group, and $S_i[t] = 1$ corresponds to joining a non future-banned group) is given by

$$\begin{aligned}\mathcal{L}_i = \prod_{t=0}^{L-2} \{&p + q - S_i[t+1]q + 2S_i[t+1]S_i[t]q \\ &- pq + S_i[t](1 - 2p - 2q + 2pq)\}.\end{aligned} \tag{7}$$

Therefore, p and q are given by

$$\begin{aligned}\arg\max_{(q,p)} \bar{\mathcal{L}} &\equiv \frac{1}{N}\sum_{i=1}^{N} \mathcal{L}_i \\ &:= \{(q,\ p)|0 \le q \le 1\ \&\ 0 \le p \le 1\}.\end{aligned} \tag{8}$$

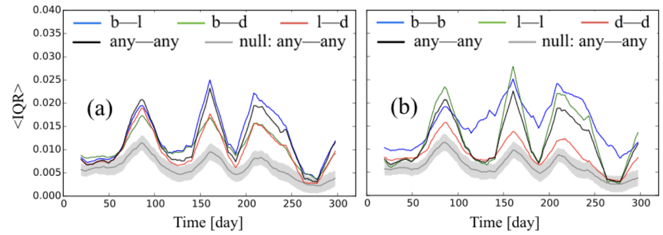

FIG. 4. (Color online) Evolution of the many-body correlation measure $\langle\mathrm{IQR}\rangle$ (i.e., average IQR) between different subpopulations of individuals. As in Fig. 3(a), 'b' means future banned, 'l' means future latent, 'd' means future self-deleting, and 'any' means irrespective of individual type. For example, 'b—d' means between future banned and future latent individuals. One standard deviation error bands are shown around the null model result. In principle, each empirical curve has its own null model, obtained by randomizing the timestamps t of the time series $s_i(t)$ (and $s_j(t)$). Since these null model results are all very similar, we only show the result for the 'any—any' case.

When doing the simulation for a given $L_{\rm ban}$, we do a separate stochastic simulation of 10,000 individuals; and since there is no history in the first step, we randomly assign the initial memory for each individual. We follow the same procedure for the other two types of individuals.

Figure 3(f) shows that q, and hence the impact of memory, is most prominent for short event-time lifetimes – consistent with the conclusion for clock-time lifetimes in Fig. 2. This could have an important implication for authorities, since it suggests that among individuals who will eventually express the most extreme support and hence become banned (i.e. future-banned individuals) the ones with shorter lifetimes will exhibit more temporal correlations and hence will exhibit more predictability in their trajectories: and it is precisely these rapidly-developing individuals who likely carry the highest risk of committing future acts.

C. Many-Body Correlations

Having characterized and quantified the trajectories of individuals, and established the increasing importance of temporal correlations at short lifetimes in both clock-time and event-time, we move to examine many-body correlations, i.e. the correlations between individuals. Though a full theory generalizing the expansion in Fig. 3(a) awaits future development, Fig. 4 shows the surprising strength and complexity of correlations that evolve over time in the system. Specifically, it shows the average information quality ratio (IQR) [28] of the group joining and leaving events for pairs of individuals of a given type, where

$$\mathrm{IQR}(X_i; X_j) = \frac{I(X_i; X_j)}{H(X_i; X_j)}. \tag{9}$$

Here X_i and X_j are two random variables measured simultaneously, $I(X_i;X_j)$ is the mutual information of the two random variables, and $H(X_i,X_j)$ is their joint entropy. IQR is normalized and lies within [0,1]. A higher value means stronger correlation: e.g. IQR = 1 means there is a one-to-one correspondence between X_i and X_j; IQR = 0 means knowing X_i says nothing about X_j. In our case, X_i and X_j represent the behaviors of individuals i and j when measured simultaneously (i.e. on the same day), therefore IQR becomes an effective measure of particle-particle correlations. More specifically, X_i and X_j are given by the signs (s_i and s_j) of the net change of the number of group memberships of user i and j on the same day, respectively.

We now describe the calculation of the average IQR (i.e. $\langle \text{IQR} \rangle$) on day t between the future banned and the future self-deleting users (denoted as b—d) in Fig. 4(a). First, we pick an individual i from the sub-population of future banned individuals, and an individual j from the sub-population of future self-deleting individuals, whose signs of the net change of the number of group memberships on day t' are $s_i(t')$ and $s_j(t')$ ($s_i(t')$ & $s_j(t') \in \{-1,0,1\}$), respectively. For $\langle \text{IQR} \rangle$ of day t, the statistics are calculated from the $(t-10)$'th day to the $(t+9)$'th day (i.e. a moving window of size 20 days). Hence, the joint probability distribution for day t is given by $P_{X_i,X_j}(x_i,x_j) = \sum_{t'=t-10}^{t+9}[\delta_{s_i(t'),x_i}\delta_{s_j(t'),x_j}]/20$, from which the marginal probability distributions $P_{X_i}(x_i)$ and $P_{X_j}(x_j)$ can be easily calculated; therefore the mutual information is given by

$$I(X_i;X_j) = \sum_{x_i,x_j} P_{X_i,X_j}(x_i,x_j)\log_2\left[\frac{P_{X_i,X_j}(x_i,x_j)}{P_{X_i}(x_i)P_{X_j}(x_j)}\right],$$

the joint entropy is given by

$$H(X_i;X_j) = -\sum_{x_i,x_j} P_{X_i,X_j}(x_i,x_j)\log_2[P_{X_i,X_j}(x_i,x_j)],$$

and eventually, the IQR between i and j is given by

$$\text{IQR}(X_i;X_j) = \frac{I(X_i;X_j)}{H(X_i;X_j)}. \tag{10}$$

We ignore pairs of users whose joint entropy is zero since they represent mostly trivial cases in which no joining/leaving events occurred. Next, we average over all combinations of the user pairs to obtain ($\langle \text{IQR} \rangle$ on day t. Since our dataset is so large, we sampled 2000 users for each individual type 10 times to obtain the mean values and their standard deviations. In addition, to improve the display we smoothened the curve by averaging over every 10 days when plotting the curves. We have investigated the effect of changing these choices of time windows and find that our main results and conclusions are unchanged.

The resulting average IQR values between trajectories from different sub-populations (Fig. 4(a)) and within the same sub-population (Fig. 4(b)) are all stronger than expected from a null model in which the order of the timestamps for x_i and x_j is randomly shuffled. In addition, the average IQRs peak around days when there are more group creation and banning events than usual. The many-body correlations between future banned users ('b—b' in Fig. 4(b)) are typically the strongest during most of the days and exceed the null model by many standard deviations, suggesting that individuals who will go on to develop the most extreme forms of support are the most synchronized. It also suggests a new dynamical collective phenomenon by which a relatively small subset of individuals manage to develop coordination within a much larger reservoir of individuals. By contrast, the correlations between future self-deleting individuals are comparable to the null model result, suggesting that the movement toward deciding to self-delete is mostly a personal one.

III. DISCUSSION AND CONCLUSION

In summary, we have identified statistical universalities in the trajectories of individuals wandering through an online extremist space, despite the heterogeneity in individuals' behaviors and final outcomes. Our findings establish the importance of temporal correlations at short lifetimes in both clock-time and event-time. These may help provide theoretical insights for impeding the development and spread of extremism online through the development and deployment of a new generation of social-media algorithms (see for example, the preliminary study in Ref. [29]). Our data and results may also help open the path toward a fuller many-body theory of human behavior [17–20] in which single-particle propagators (individuals) successively scatter through dynamical groups that themselves comprise other single-particle propagators, thereby yielding a coupled hierarchy of propagators in a full diagrammatic expansion [26].

Assuming that terrorist events emerge from a complex chain of behaviors, such a real-world chain would undoubtedly include many complex steps such as adopting an extremist ideology, thinking about engaging in violence, acquiring the necessary materials and/or training, and finally committing the offense [2], while the Internet facilitates the development of this chain of behaviors [3–5]. Previous studies have primarily focused on statistical analysis of the correlations between various factors and the off-line terror attacks, with the aim of identifying potential extremists at an early stage [2–5]. This identification faces the challenge that there may be no simple set of factors that can effectively distinguish high-risk individuals from lower-risk ones[1–3], and also that each individual's development toward radical ideology and extremist behavior is likely highly dynamic. As a result, an extremist can emerge from a group that may previously have been identified as very unlikely to produce extremists. In this work, we adopted a different approach by looking for universal patterns and principles of human behaviors in a typical online extremist network and modeling them with classical physics approaches. As Gill et al. point out in Ref. 5, there is an important cautionary statement that violent radicalization should be framed as cyber-enabled rather than as cyber-dependent. Moreover, enabling factors differ from case to case depending on need and circumstance. Despite this, we hope that the present paper will stimulate new collaborations at the interface between emergent quantitative modeling and existing expertise in case-study approaches. While mitigating the risk of future events is certainly a stretch, our work might conceivably be further developed to help indicate which individuals show radical support online and identify others who do, too. This in turn could help devise strategies to better counter the spread of radical propaganda online. We also are aware that the vast majority of individuals who are flagged as showing radical support will never actually commit any illegal act. However, knowing that they do can help inform the reactions of security forces, together with the contextual factors and specific circumstances of every individual, as highlighted by studies such as the one of Gill et al. in Ref. 5.

Acknowledgments We thank A. Gabriel, A. Kuz, J. Nearing and T. Curtright for initial help with data and discussions. NFJ acknowledges funding under National Science Foundation (NSF) grant CNS 1522693 and Air Force (AFOSR) grant FA9550-16-1-0247. The views and conclusions contained herein are solely those of the authors and do not represent official policies or endorsements by any of the entities named in this paper.

[1] P. Gill, Lone Actor Terrorists: A Behavioral Analysis. (Routledge, London, 2015)
[2] P. Gill, J. Horgan, P. Deckert. Bombing alone: Tracing the motivations and antecedent behaviors of lone-actor terrorists. Journal of Forensic Sciences **59**, 425 (2014)
[3] P. Gill, E. Corner. Lone-actor terrorist use of the Internet and behavioural correlates. In Terrorism Online: Politics, Law, Technology and Unconventional Violence. Eds. L. Jarvis, S. Macdonald and T.M. Chen (Routledge, London, 2015)
[4] L. Von Behr, A. Reding, C. Edwards, L. Gribbon. Radicalization in the Digital Era: The Use of the Internet in 15 Cases of Terrorism and Extremism. (RAND, Santa Monica CA, 2013)
[5] P. Gill, E. Corner, M. Conway, A. Thornton, M. Bloom, J. Horgan. Terrorist Use of the Internet by the Numbers Quantifying Behaviors, Patterns, and Processes. Criminology and Public Policy **16**, 99 (2017)
[6] BBC News. Russian students targeted as recruits by Islamic State. http://linkis.com/dmFgu (July 24, 2015)
[7] BBC News. IS group to step up attacks on Europe. http://www.bbc.com/news/world-europe-38179653 (Dec. 2, 2016)
[8] J.M. Berger, H. Perez. The Islamic State's diminishing returns on Twitter. GW Program on Extremism Vol. 2 (2016)
[9] N.F. Johnson, M. Zheng, Y. Vorobyeva, A. Gabriel, H. Qi, N. Velasquez, P. Manrique, D. Johnson, E. Restrepo, C. Song, S. Wuchty, S. New online ecology of adversarial aggregates: ISIS and beyond. Science **352**, 1459 (2016)
[10] J. Pagliery. Suspect in congressional shooting was Bernie Sanders supporter, strongly anti-Trump. http://www.cnn.com/2017/06/14/homepage2/james-hodgkinson-profile/index.html (June 15, 2017)
[11] H. Yan, D. Simon, A. Graef. Campus killing: Suspect is a member of 'Alt-Reich' Facebook group. http://www.cnn.com/2017/05/22/us/university-of-maryland-stabbing/index.html (May 22, 2017)
[12] R. Axelrod. The dissemination of culture. J. Confl. Resolut. **41**, 203 (1997)
[13] D. Centola, J.C. Gonzalez-Avella, V.M. Eguiluz, M. San Miguel. Homophily, Cultural drift, and the co-evolution of cultural groups. J. Confl. Resolut. **51**, 905 (2007)
[14] T.T. Zin, P. Tin and T. Toriu. A series of stochastic models for human behavior analysis. 2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Seoul, 2012, pp. 3251-3256. doi: 10.1109/ICSMC.2012.6378292
[15] Social Phenomena: From Data Analysis to Models. Eds: B. Goncalves and N. Perra (Springer, New York, 2015)
[16] A.L. Barabasi. The origin of bursts and heavy tails in human dynamics. Nature **435**, 207 (2005)
[17] M.C. Gonzalez, C.A. Hidalgo, A.L. Barabasi. Understanding individual human mobility patterns. Nature **453**, 779 (2008)
[18] G.M. Viswanathan, G.E. da Luz, E.P. Raposo, H.E. Stanley. The Physics of Foraging. (Cambridge University Press, Cambridge, 2011)
[19] I.D. Couzin, J. Krause, N.R. Franks, S.A. Levin. Effective leadership and decision-making in animal groups on the move. Nature **433**, 513 (2005)
[20] S. Gavrilets. Collective action and the collaborative brain. J. R. Soc. Interface **12**, 20141067 (2015)
[21] A. Molini et al. First passage time statistics of Brownian motion with purely time dependent drift and diffusion. Physica A: Statistical Mechanics and its Applications. **390**, 1841 (2011)
[22] R.N. Mantegna, H.E. Stanley. Introduction to Econophysics: Correlations and Complexity in Finance (Cambridge University Press, New York, 2007)
[23] B.G. Malkiel. A Random Walk Down Wall Street (Norton, New York, 2003)
[24] N.F. Johnson, P. Jefferies, P.M. Hui. Financial Market Complexity: What Physics Can Tell Us About Market Behavior (Oxford University Press, Oxford, 2003)
[25] A.L. Barabasi, H.E. Stanley. Fractal Concepts in Surface Growth (Cambridge University Press, Cambridge, 1995)
[26] R.D. Mattuck. A Guide to Feynman Diagrams in the Many-Body Problem (Dover, New York, 1992)
[27] A. Clauset, C.R. Shalizi, and M.E.J. Newman. Power-law distributions in empirical data. SIAM Review **51**, 661 (2009)

[28] D. R. Wijaya, R. Sarno, and E. Zulaika. Information Quality Ratio as a novel metric for mother wavelet selection. Chemometrics and Intelligent Laboratory Systems **160**, 59 (2017)

[29] N.F. Johnson et al. Population polarization dynamics and next-generation social media algorithms. arXiv preprint arXiv:1712.06009 (2017)