Aalto University

Choi, Jeehye; Hiraoka, Takayuki; Jo, Hang Hyun

Individual-driven versus interaction-driven burstiness in human dynamics

Individual-driven versus interaction-driven burstiness in human dynamics: The case of Wikipedia edit history

Jeehye Choi [1,2,*] Takayuki Hiraoka [3] and Hang-Hyun Jo [2,†]

[1]*Asia Pacific Center for Theoretical Physics, Pohang 37673, Republic of Korea*
[2]*Department of Physics, The Catholic University of Korea, Bucheon 14662, Republic of Korea*
[3]*Department of Computer Science, Aalto University, Espoo FI-00076, Finland*

The origin of non-Poissonian or bursty temporal patterns observed in various data sets for human social dynamics has been extensively studied, yet its understanding still remains incomplete. Considering the fact that humans are social beings, a fundamental question arises: Is the bursty human dynamics dominated by individual characteristics or by interaction between individuals? In this paper we address this question by analyzing the Wikipedia edit history to see how spontaneous individual editors are in initiating bursty periods of editing, i.e., individual-driven burstiness, and to what extent such editors' behaviors are driven by interaction with other editors in those periods, i.e., interaction-driven burstiness. We quantify the degree of initiative (DoI) of an editor of interest in each Wikipedia article by using the statistics of bursty periods containing the editor's edits. The integrated value of the DoI over all relevant timescales reveals which is dominant between individual-driven and interaction-driven burstiness. We empirically find that this value tends to be larger for weaker temporal correlations in the editor's editing behavior and/or stronger editorial correlations. These empirical findings are successfully confirmed by deriving an analytic form of the DoI from a model capturing the essential features of the edit sequence. Thus our approach provides a deeper insight into the origin and underlying mechanisms of bursts in human social dynamics.

I. INTRODUCTION

Since the seminal paper by Barabási [1], the origin of bursty temporal patterns in human dynamics has been extensively debated for the last few decades (see the review book [2] and references therein). Here the bursts indicate the rapidly occurring events in short time periods that are separated by long periods of low activity. In his paper, Barabási argued that non-Poissonian or bursty patterns observed in the email communication data set can be understood by considering how individuals manage the incoming emails according to their priorities. In contrast, Malmgren *et al.* [3] suggested that bursty patterns in the same email data set can be largely explained by cyclic behaviors such as daily and weekly cycles of humans. Later Jo *et al.* [4] showed that deseasoning daily and weekly cycles from the time series of mobile phone communication cannot entirely remove the burstiness in the time series, implying that the remaining burstiness might be due to various other factors affecting the human dynamics.

Note that the above studies have focused on the individual time series. However, considering the fact that humans are social beings, the effects of social interaction on the bursty human dynamics should not be ignored. Here a fundamental question arises: Are bursts in human dynamics more likely to be the consequences of intrinsic bursty characteristics of individuals or more driven by the interaction between individuals? Borrowing the terms in the network science [5–7] we refer to this question as *node burstiness versus link burstiness*. Karsai *et al.* [8] addressed a similar question by analyzing the time series of individuals (nodes) and their ties (links) in the mobile phone communication data set. They found that the bursty behavior of nodes is dominated by that of links incident to the nodes, hence concluded that "burstiness is a property of the links rather than of the nodes." Here we remark that the mobile phone communication data set consists of interaction events only. In particular, each mobile phone call can be described by a tuple (i, j, t) in that a caller i makes a call to a receiver j at time t [8–11]. Therefore the mobile phone communication data set has also been studied in the framework of temporal networks [12–14], because the temporal network can be defined as a set of interaction events or tuples (i, j, t). We note that a temporal network model has recently been suggested to reconcile node burstiness and link burstiness [15] rather than to contrast one with the other. However, such a temporal-network approach may systematically preclude individual activities that are not described by tuples (i, j, t), such as watching a movie alone or posting to a personal blog spontaneously. These events that do not

*Present address: Research Institute for Nanoscale Science and Technology, Chungbuk National University, Cheongju 28644, Republic of Korea.
†h2jo@catholic.ac.kr

necessarily imply interaction between individuals are referred to as stand-alone events and denoted by (i, t); they can also be important in understanding the origin of bursty human dynamics.

As for an illustrative case study regarding the origin of bursty human dynamics, one can analyze the edit history of the self-organized online collaborative encyclopedia, Wikipedia [16,17], among others such as the online forum data set [18] and Twitter data set [19]. It is well known that the temporal patterns of edit sequences of Wikipedia articles as well as those of editors are bursty [17,20–27]. Unlike mobile phone communication, in which every event is a call or a message between users, editing Wikipedia articles is not evidently a communicational act between editors and thus could be described as a collection of stand-alone events. Nevertheless, it has been repeatedly shown that interaction between editors does exist and in fact plays an important role in revising articles. Such interaction can be inferred, e.g., by comparing different versions (or revisions) of the article that were modified by different editors [17,28–31], by considering the orders of editors who edited the same article [32–35], and by relating all the active editors who edited the same article to each other [36]. Other works considered the interaction between editors not only in articles but also in their talk pages [37,38]. Using the definition of the pairwise interaction between editors, social networks of editors were constructed and analyzed [28,32,35,36]. Finally, bipartite networks between articles and editors have also been studied [29,39–41].

In this paper, by analyzing the edit sequences of Wikipedia articles we aim to understand the role of individuals in the bursty collective dynamics by looking at how spontaneous individual editors are in initiating bursty periods of editing and to what extent such editors' behaviors are driven by interaction with other editors in those periods. These questions can be formulated in terms of *individual-driven burstiness versus interaction-driven burstiness*, which can be seen as an elaboration of the issue on node burstiness versus link burstiness. To infer the interaction between editors we detect bursty periods for a given timescale [42] from the edit sequence of each Wikipedia article. Here the editors who edited the same article during the same bursty period are assumed to have interacted with each other, irrespective of the contents modified by editors. Our approach is similar to that by Karsai *et al.* [8] as bursty periods are detected from a sequence of calls or edits, but different in the sense that we focus on the information on who initiates the bursty periods, i.e., who makes the first edit of the bursty period.

As human activities are sometimes spontaneous and other times interaction-driven, the individual human behavior might not be understood only by one of individual-driven burstiness and interaction-driven burstiness. Therefore, we take an approach of quantifying the degree of initiative (DoI) of the editor of interest in a given article using the statistics of bursty periods containing the editor's edits. Since the bursty periods are detected for a given timescale [42], the DoI is also a function of the timescale. By scanning the entire range of the timescale, we obtain the DoI curve, from which the area under the curve (AUC) is calculated. A large value of the AUC can be interpreted as the dominance of individual-driven burstiness over interaction-driven burstiness and vice versa.

To investigate features in the edit sequences that are relevant in understanding the observed AUC values, we correlate the AUC value with several measures for temporal and editorial correlations. Finally, we confirm the empirical findings by devising and analyzing a model capturing the essential features of edit sequences.

The paper is organized as follows. In Sec. II we describe our approach for analyzing the edit sequences of Wikipedia articles by means of bursty periods, the degree of initiative, and the area under the curve. In Sec. III we devise and analyze a model capturing the essential features of the edit sequences to confirm the empirical findings in Sec. II. Finally, we conclude our work with some remarks on future works in Sec. IV.

II. EMPIRICAL ANALYSIS

A. Data set

To scrutinize the issue on individual-driven burstiness versus interaction-driven burstiness for the origin of bursts in human dynamics, we analyze a data set of the English Wikipedia dump on October 2, 2015 [43]. The data set originally contains 11 994 178 general articles, which we call articles in our paper. Among them, we analyze articles that have not been merged with other articles, the number of which is 4 978 964. The number of editors who contributed those articles is 40 057 921. The preprocessed data for each article contain the temporal and editorial information of edits; each edit is recorded with the editor ID and the time stamp in the resolution of seconds. See the sample of the data in Figs. 1(a) and 1(b). The data analyzed in our work are available in the public repository [44]. We remark that we do not take any other information on edits, such as the size and content of edits, into account for the analysis.

B. Degree of initiative

An edit sequence of each Wikipedia article is given as a time-ordered set of edits. For the article with n edits, the ith edit for $i = 1, \ldots, n$ is associated with the editor c_i and timing t_i. Among the editors of the article, a particular editor is chosen to be called *ego*, while all other editors are called *alters*. For simplicity we do not distinguish alters; each edit is made either by the ego or by the alter, i.e., $c_i \in \{E, A\}$, where E and A stand for the ego and the alter, respectively. Note that from the sequence of edit timings, i.e., $\{t_i\}_{i=1,\ldots,n}$, one obtains the sequence of inter-edit times (IETs) by the definition of $\tau_i \equiv t_i - t_{i-1}$.

To define the interaction between editors, we assume that editors who edited the same article within the same chunk of time period have interacted with each other. Precisely, we adopt the notion of bursty trains or bursts [2,42]: For a given timescale Δt, a burst is defined as a set of consecutive edits such that the IET between any two consecutive edits in the burst is smaller than or equal to Δt, while edits in different bursts are separated by IETs larger than Δt. See Fig. 1(c) for an illustrative example. Then we categorize the detected bursts, whose number is denoted by m, into three groups according to the editorial information: (i) A burst of which the first edit is made by the ego is called an *initiated burst*. (ii) A burst that is not initiated by the ego but contains
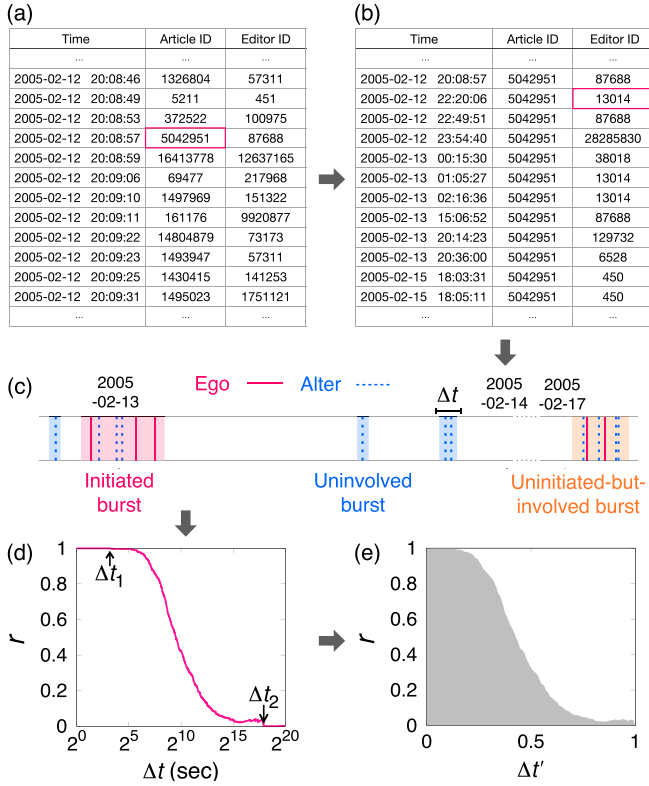
FIG. 1. Diagram for the data analysis procedure. (a) Part of the preprocessed data of Wikipedia edit history. Each line contains the time stamp, article ID, and editor ID of the edit. (b) Part of the edit sequence of a given article (e.g., article 5 042 951) containing the time stamp and the editor ID per edit. (c) Visualization of part of the edit sequence for article 5 042 951. Each edit is denoted by a vertical line in the horizontal time axis. Among editors who edited the article, we choose a particular editor or ego (e.g., editor 13 014), while all other editors are called alters. Edits by the ego and alters are denoted by red solid lines and blue dotted lines, respectively. For a given timescale Δt, bursts are detected and categorized into three groups: initiated (red area), uninitiated-but-involved (orange area), and uninvolved bursts (blue area). (d) The degree of initiative (DoI), r in Eq. (1), for editor 13 014 in article 5 042 951, with Δt in seconds. See the text for the definitions of Δt_1 and Δt_2. (e) The DoI curve as a function of the normalized timescale $\Delta t'$ in Eq. (2), from which the area under the DoI curve is calculated as ≈ 0.45.

ego's edits is called an *uninitiated-but-involved burst*. (iii) A burst without ego's edits is called an *uninvolved burst*. The numbers of initiated bursts, uninitiated-but-involved bursts, and uninvolved bursts are respectively denoted by m_{init}, m_{inv}, and m_{other}, satisfying $m = m_{\text{init}} + m_{\text{inv}} + m_{\text{other}}$. Using the first two numbers we define the degree of initiative (DoI) for the ego in the given article as follows:

$$r \equiv \frac{m_{\text{init}}}{m_{\text{init}} + m_{\text{inv}}}, \qquad (1)$$

which essentially quantifies how often the ego initiates bursts while being active in editing.

For a fixed timescale Δt, a larger value of r implies more initiative behavior of the ego at that timescale. However, it is not obvious which timescale is the most relevant for understanding the initiative behavior. Therefore, we study the DoI

for the entire range of the timescale for systemic investigation. If Δt is smaller than the minimum IET of the edit sequence, each edit constitutes a burst of size one on its own, implying that $m_{\text{inv}} = 0$, hence $r = 1$. As Δt increases, bursts are merged with each other, and the total number of bursts, m, decreases. In particular, m_{init} either decreases or remains the same, but never increases, which is the main driving force for the overall decreasing r. The increasing behavior of r is found only when m_{inv} decreases. Precisely, m_{inv} decreases by one whenever an uninitiated-but-involved burst is merged either with the preceding initiated burst or with another uninitiated-but-involved burst. However, if the number of ego's edits is much smaller than the total number of edits, as is the case in our work, it is more common to find a merger of an uninvolved burst and its following initiated burst into an uninitiated-but-involved burst, leading to the overall increasing m_{inv}, hence the overall decreasing r. Finally, if Δt is equal to or larger than the maximum IET in the edit sequence, all edits belong to a single burst. This burst is either an initiated burst or an uninitiated-but-involved burst, ending up with $r = 1$ or 0, respectively, depending on who made the first edit to the article. Such dependence on the initial condition is somewhat excessive and makes the further analysis less robust. For the edit sequences where the ego makes the first edit, to ensure $r = 0$ for Δt equal to or larger than the maximum IET in the edit sequence, we add a dummy edit by an alter before the beginning of the edit sequence; we suppose the zeroth edit made by an alter $c_0 = A$ at time $t_0 = t_1 - \tau_1$ for some IET τ_1 (to be discussed).

To guarantee the statistical significance of the analysis, we first choose articles with more than 10^4 edits from the Wikipedia data set, leaving us with 697 articles. For each article we choose human editors who have made at least 1% of edits to the article [45]. In total there are 3099 such editors, each of whom is designated as the ego. As a result we are left with 4634 article-ego pairs to be analyzed. As an example, we plot in Fig. 1(d) the DoI curve for editor 13 014, who edited 559 times the article 5 042 951 with $n = 20\,683$. As expected, the DoI curve overall decreases with the timescale Δt.

In each of the DoI curves obtained from the data set, we observe two characteristic timescales: At the first timescale, denoted by Δt_1, the DoI curve starts to deviate from $r = 1$ as uninitiated-but-involved bursts appear, i.e., $m_{\text{inv}} > 0$. At the second timescale, denoted by Δt_2, the DoI curve reaches zero as all initiated bursts disappear, i.e., $m_{\text{init}} = 0$. The timescale Δt_1 is equal to the minimum IET between the ego's edit and its preceding edit by the alter, provided that the alter's edit is not part of an initiated burst detected at the timescale Δt_1. This may indicate a tendency that the ego makes edits only after Δt_1 since the edits by the alters. If editing by different editors in the same burst may imply the interaction between those editors, Δt_1 could be interpreted as the minimal timescale for the ego's interaction with other editors. Similarly, the timescale Δt_2 can be a good proxy for the maximum IET between the ego's edit and its preceding edit by the alter. That is, the ego tends to make edits within Δt_2 since the edits by the alters, implying that Δt_2 could be interpreted as the maximal timescale for the ego's interaction with other editors. We find in Fig. 2(a) that the population of editors can be well described by some typical values of Δt_1 and Δt_2. Here Δt_1 turns out to be shorter than the order of one day
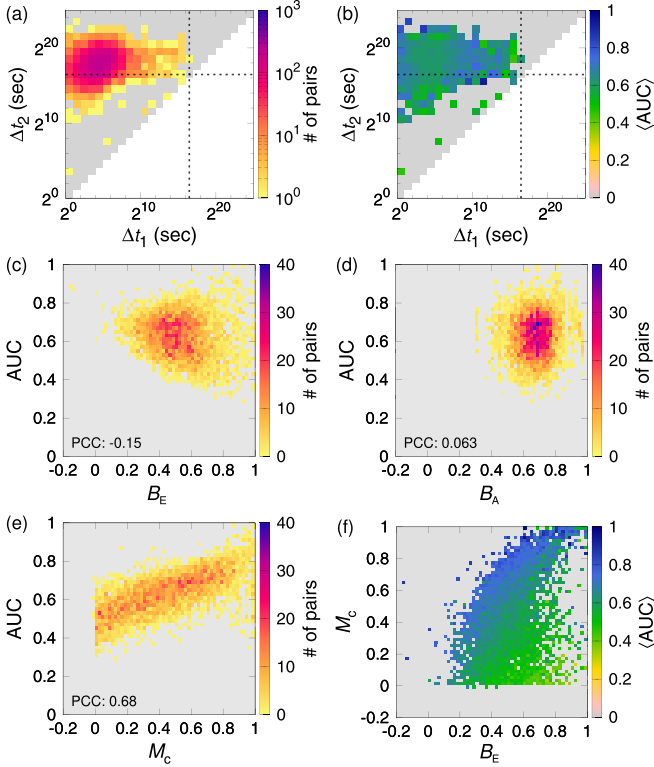
FIG. 2. (a) Heat map of the number of article-ego pairs having the same characteristic timescales Δt_1 and Δt_2. (b) Heat map of the averaged value of the area under the DoI curve, denoted by $\langle \mathrm{AUC} \rangle$, for the pairs having the same characteristic timescales Δt_1 and Δt_2. In panels (a) and (b), log-binned Δt_1 and Δt_2 have been used for the calculation. Both horizontal and vertical dashed lines denote one day. (c)–(e) Heat maps of the number of article-ego pairs in the spaces of (B_{E}, AUC) (c), (B_{A}, AUC) (d), and (M_c, AUC) (e). See the text for the definitions of B_{E}, B_{A}, and M_c. In each of panels (c)–(e), we also present the value of the Pearson correlation coefficient (PCC). (f) Heat map of $\langle \mathrm{AUC} \rangle$ for the pairs having the same values of B_{E} and M_c. In all panels the gray color means no data in the area.

in almost all article-ego pairs, possibly due to the fact that we have chosen article-ego pairs with relatively active editors. We also find 14 article-ego pairs (for three distinct egos) with $\Delta t_1 = \Delta t_2$, because these pairs have the editorial structure of $\mathrm{A} \cdots \mathrm{AE} \cdots \mathrm{EA} \cdots \mathrm{A}$. This implies that r drops from 1 to 0 at the timescale $\Delta t_1 \, (= \Delta t_2)$; hence these pairs will be excluded for the further analysis in the next subsection.

Finally, we remark that the number of article-ego pairs to which we have added the dummy zeroth edit by the alter is quite small (14 among 4634 pairs) and that the IET between the zeroth and first edits, τ_1, has been set to be Δt_1 not to affect the empirical results for Δt_1 and Δt_2.

C. Area under the DoI curve

To fully characterize the DoI curves, we also look at the overall decreasing behavior of each of those curves in the range of $[\Delta t_1, \Delta t_2]$. From the empirical results of the DoI curves, we find various decreasing patterns such as in convex, concave, and linear manners. To quantify such diverse behaviors by a single value, we calculate the area under the

DoI curve (AUC) in a normalized range of the timescale. For $\Delta t \in [\Delta t_1, \Delta t_2]$, we define a normalized timescale as

$$\Delta t' \equiv \frac{\log \Delta t - \log \Delta t_1}{\log \Delta t_2 - \log \Delta t_1}. \qquad (2)$$

Note that $0 \leqslant \Delta t' \leqslant 1$. The DoI curve as a function of $\Delta t'$ is used to calculate the AUC, e.g., as shown in Fig. 1(e). The value of AUC represents the overall initiative behavior of the ego for the entire range of relevant timescales. Using the AUC one can also compare different article-ego pairs more conveniently irrespective of their characteristic timescales of Δt_1 and Δt_2. In sum, each DoI curve for the article-ego pair can be characterized in terms of three quantities, namely, Δt_1, Δt_2, and the AUC.

Empirical results of Δt_1, Δt_2, and the AUC obtained for 4620 article-ego pairs are summarized in Fig. 2(b), where we take the average of AUC values, denoted by $\langle \mathrm{AUC} \rangle$, for the set of article-ego pairs having the same values of Δt_1 and Δt_2, and plot $\langle \mathrm{AUC} \rangle$ in the space of $(\Delta t_1, \Delta t_2)$. The value of $\langle \mathrm{AUC} \rangle$ turns out to be overall independent of Δt_1 and Δt_2.

We investigate features in the edit sequences that are relevant in understanding the observed AUC values. For this, we characterize the edit sequences by two types of features, i.e., temporal correlations and editorial correlations. As for the features of the first type, we consider the set of IETs followed by the ego's edits, i.e., $\mathcal{T}_{\mathrm{E}} \equiv \{\tau_i | c_i = \mathrm{E}\}$, as well as the set of IETs followed by alters' edits, i.e., $\mathcal{T}_{\mathrm{A}} \equiv \{\tau_i | c_i = \mathrm{A}\}$. Each IET in $\mathcal{T}_{\mathrm{E}}$ indicates how long the ego waits for the next edit since the latest edit, whether it was made by the ego or by the alter. Thus heterogeneity of IETs in $\mathcal{T}_{\mathrm{E}}$ can reveal the temporal property of the ego's editing behavior. In general, heterogeneity of inter-event times in event sequences has been extensively discussed to show the temporal correlations in various empirical time series [2]. To quantify the heterogeneity of IETs in $\mathcal{T}_{\mathrm{E}}$, we adopt the burstiness measure suggested for finite event sequences [46]:

$$B_{\mathrm{E}} \equiv \frac{\sqrt{n_{\mathrm{E}} + 1} \, \sigma_{\mathrm{E}} - \sqrt{n_{\mathrm{E}} - 1} \, \mu_{\mathrm{E}}}{(\sqrt{n_{\mathrm{E}} + 1} - 2) \sigma_{\mathrm{E}} + \sqrt{n_{\mathrm{E}} - 1} \, \mu_{\mathrm{E}}}, \qquad (3)$$

where n_{E} is the number of IETs in $\mathcal{T}_{\mathrm{E}}$. σ_{E} and μ_{E} denote the standard deviation and average of IETs in $\mathcal{T}_{\mathrm{E}}$, respectively. The range of B_{E} is $[-1, 1]$. The positive B_{E} indicates a bursty temporal pattern in the ego's editing behavior, while $B_{\mathrm{E}} = 0$ if the ego's editing behavior can be described by a Poisson process. The negative B_{E} is observed for relatively regular temporal patterns. We also consider the burstiness measure for the alters obtained from $\mathcal{T}_{\mathrm{A}}$, denoted by B_{A}, in a similar manner.

The second type of feature we consider is the correlation between two consecutive editors in the editor sequence, which can be quantified in terms of the Pearson correlation coefficient (PCC):

$$M_c \equiv \frac{1}{n-1} \sum_{i=1}^{n-1} \frac{[h(c_i) - \mu_1][h(c_{i+1}) - \mu_2]}{\sigma_1 \sigma_2}, \qquad (4)$$

where the editorial information is transformed to the numerical value by the function h, namely, $h(c_i = \mathrm{E}) = 1$ and $h(c_i = \mathrm{A}) = 0$. μ_1 (μ_2) and σ_1 (σ_2) denote the average and standard deviation of h values except for the last (the first)

FIG. 3. (a) Part of time series for the article-ego pair that is characterized by small B_E and large M_c (hence, individual-driven burstiness). (b) Part of time series for the article-ego pair that is characterized by large B_E and small M_c (hence, interaction-driven burstiness). Red (blue) vertical lines denote the ego's (the alter's) edits. Time intervals have been chosen to contain 20–30 edits by the ego in both cases.

editor, respectively. Positive M_c implies a tendency of the ego's edit (alter's edit) to be followed by the ego's edit (alter's edit). The opposite tendency is implied by the negative M_c, while $M_c = 0$ indicates the absence of correlations between two consecutive editors.

For each article-ego pair we calculate the values of B_E, B_A, and M_c to correlate them with the AUC value for the pair. Figures 2(c)–2(e) show heat maps of the number of article-ego pairs in the spaces of (B_E, AUC), (B_A, AUC), and (M_c, AUC), respectively. Then for each case, we calculate the value of PCC ρ: We observe (i) a negative correlation between B_E and AUC values ($\rho = -0.15$), (ii) a slightly positive correlation between B_A and AUC values ($\rho = 0.063$), and (iii) a strongly positive correlation between M_c and AUC values ($\rho = 0.68$).

Observation (i) implies the tendency that the stronger temporal correlation in the ego's editing behavior (i.e., larger B_E) leads to the smaller AUC values. This can be intuitively understood by the following argument: The larger burstiness measure indicates a more heterogeneous IET distribution, typically with a higher peak at small IETs and a heavier tail at large IETs. As Δt increases, small IETs followed by the ego's edits enhance the merger of initiated bursts and their preceding bursts, resulting in the fast-decaying m_{init} and DoI curve. On the other hand, for the range of large Δt, large IETs followed by the ego's edits slow down the merger of initiated bursts and their preceding bursts, resulting in the slow decaying m_{init} and DoI curve. Combining these two effects, one gets the initially fast-decaying and then slow-decaying DoI curve, hence the overall convex DoI curve. The more convex a curve is, the smaller the AUC value is calculated from it. This argument explains the observed negative correlation between B_E and AUC values.

Observation (ii), i.e., the slightly positive correlation between B_A and AUC values, can be explained similarly but in an opposite manner to the case with observation (i). More importantly, we find that the effect of B_A on the AUC is not as significant as that of B_E on the AUC. It is probably because the majority of edits in the article are made by the alters, implying that most of IETs followed by the alters' edits have only marginal effects on the statistics of m_{init} and m_{inv}.

Finally, observation (iii), i.e., a strongly positive correlation between M_c and AUC values, allows us to argue that as the ego's edits become more segregated from the alters' edits due

to stronger editorial correlation (i.e., larger M_c), the number of uninitiated-but-involved bursts, m_{inv}, would be smaller than otherwise. The overall smaller m_{inv} leads to the more elevated DoI curve [see Eq. (1)], hence the larger AUC value.

Let us briefly discuss implications of the empirical results. For the sake of simplicity, we assume that the temporal properties of the alters' edits, characterized by B_A, are given or fixed. This assumption allows us to focus on the effects of the ego's temporal pattern and the editorial correlation on the AUC value. On the one hand, a large AUC value would be achieved for a small B_E and large M_c, corresponding to the top left area in Fig. 2(f). Let us consider an edit sequence where the ego's edits are well segregated from those of alters (i.e., large M_c), while the ego's temporal pattern is largely homogeneous ($0 \lesssim B_E \ll 1$). This implies that the ego is rarely affected by interaction with other editors and more likely to be described by the individual behavior. Therefore, the ego's editorially correlated but temporarily random behavior can be interpreted as the individual-driven burstiness. See Fig. 3(a) for an exemplary time series corresponding to the case dominated by the individual-driven burstiness.

On the other hand, a small AUC value can be found for a large B_E and small M_c, corresponding to the bottom right area in Fig. 2(f). One can consider an edit sequence where the ego's edits are well mixed with the alters' edits (small M_c), and the IETs between the ego's edits and their preceding edits by alters are mostly small but often very large (large B_E). Therefore, the ego seems to interact with other editors so that the ego's such behavior can be described by the interaction-driven burstiness. See Fig. 3(b) for an exemplary time series corresponding to the case dominated by the interaction-driven burstiness.

In conclusion, by correlating the observed AUC value with the burstiness measures for the IETs followed by the ego and by the alters and the Pearson correlation coefficient between two consecutive editors, we find that weaker (stronger) temporal correlations in the ego's (alters') editing behavior and/or stronger editorial correlations tend to result in the larger values of AUC. These empirical findings are explained in terms of individual-driven burstiness and interaction-driven burstiness. For more systematic investigation of the mechanisms behind such tendencies we study a model with tunable temporal and editorial correlations in the next section.

III. MODEL

To investigate the mechanisms behind the empirical observation for the DoI curves and corresponding AUC values, we devise and analyze a model that generates edit sequences with tunable temporal and editorial correlations. For the temporal correlations, we consider the distribution of the IETs followed by the ego's edits, denoted by $P_E(\tau)$. We also consider the IET distribution for the alter, denoted by $P_A(\tau)$, in a similar manner. For the editorial correlations, we consider the conditional probability $p(c_i|c_{i-1})$, namely, $q_{E|E}$, $q_{A|E}$, $q_{E|A}$, and $q_{A|A}$. For example, $q_{E|A}$ denotes the probability that the ego's edit follows the alter's edit, irrespective of the timings of edits. These probabilities are not independent of each other but satisfy the following conditions:

$$q_{E|E} + q_{A|E} = 1, \tag{5}$$

$$q_{E|A} + q_{A|A} = 1, \tag{6}$$

$$q_{A|E}q_E = q_{E|A}q_A, \tag{7}$$

where q_E ($q_A = 1 - q_E$) denotes the fraction of the ego's edits (alter's edits) in the edit sequence. The condition in Eq. (7) indicates the asymptotic balance between the frequency of the sequence "EA" and that of "AE" in the editor sequence. As a result, we are left with two independent parameters, i.e., q_E and $q_{E|E}$.

To generate an edit sequence, we begin with an alter's edit at the initial time, i.e., $c_0 = A$ and $t_0 = 0$. Then the ith edit for $i = 1, \ldots, n$ is generated as follows:

(1) The editor of the ith edit, i.e., $c_i \in \{E, A\}$, is determined using $p(c_i|c_{i-1})$, conditioned by the editor of the $(i-1)$th edit, i.e., c_{i-1}.

(2) If $c_i = E$ (A), the preceding IET τ_i is randomly drawn from $P_E(\tau)$ [$P_A(\tau)$] to determine the timing of the ith edit as $t_i = t_{i-1} + \tau_i$.

Once the edit sequence of $n + 1$ edits is generated, it is analyzed to obtain the DoI curve and its corresponding AUC value.

We analyze our model by deriving the analytic form of the DoI, r, in terms of q_E and $q_{E|E}$ as well as the arbitrary functional forms of the IET distributions $P_E(\tau)$ and $P_A(\tau)$. Let us consider an edit sequence of $n + 1$ edits, i.e., n IETs. The distribution of these IETs, denoted by $P_{all}(\tau)$, is related to $P_E(\tau)$ and $P_A(\tau)$ as

$$P_{all}(\tau) = q_E P_E(\tau) + q_A P_A(\tau). \tag{8}$$

We denote by m the number of bursts detected from the edit sequence for a given timescale Δt. We define the fractions of bursts of different types discussed in the previous section as follows:

$$p_{init} \equiv \frac{m_{init}}{m}, \tag{9}$$

$$p_{inv} \equiv \frac{m_{inv}}{m}, \tag{10}$$

$$p_{other} \equiv \frac{m_{other}}{m}. \tag{11}$$

Then the DoI in Eq. (1) is rewritten as

$$r = \frac{p_{init}}{1 - p_{other}}. \tag{12}$$

First, we derive the analytic form of p_{init}. Since each IET larger than Δt separates two consecutive bursts detected at the timescale Δt, the number of IETs larger than Δt, denoted by n', is related to m as

$$n' = m - 1. \tag{13}$$

Further, n' can also be written in terms of the cumulative IET distribution as

$$n' = n[1 - F(\Delta t)], \tag{14}$$

where

$$F(\Delta t) \equiv \int_0^{\Delta t} P_{all}(\tau)d\tau. \tag{15}$$

Therefore, one gets for $m \gg 1$ [47,48]

$$m \simeq n[1 - F(\Delta t)]. \tag{16}$$

Each initiated burst is detected whenever an ego's edit follows any previous edit after an IET larger than Δt. Precisely,

$$m_{init} = n \Pr[c_i = E \cap \tau_i > \Delta t] = nq_E[1 - F_E(\Delta t)], \tag{17}$$

where

$$F_E(\Delta t) \equiv \int_0^{\Delta t} P_E(\tau)d\tau. \tag{18}$$

We drop Δt from now on for the sake of simplicity. From Eqs. (16) and (17) we obtain p_{init} as

$$p_{init} = \frac{q_E(1 - F_E)}{1 - F}. \tag{19}$$

Second, for the derivation of p_{other} one needs to know the fraction of bursts only consisting of the alter's edits among bursts whose first edit is made by the alter. For this, we define the fraction of bursts only with the alter's edits as

$$f \equiv \frac{m_{other}}{m_{inv} + m_{other}} \tag{20}$$

to get

$$p_{other} = (1 - p_{init})f. \tag{21}$$

Here f can be explicitly written as

$$f = \frac{\sum_{b=1}^{\infty} \Pr[c_2 = \cdots = c_b = A|c_1 = A]}{\sum_{b=1}^{\infty} \sum_{c_2 \in \{E,A\}} \cdots \sum_{c_b \in \{E,A\}} \Pr[c_2, \cdots, c_b|c_1 = A]}. \tag{22}$$

To calculate the numerator on the right-hand side of Eq. (22), we consider the probability that once an alter's edit is made, the next event is also made by the alter within an IET smaller than or equal to Δt. This probability is given by $q_{A|A}F_A$, where

$$F_A(\Delta t) \equiv \int_0^{\Delta t} P_A(\tau)d\tau. \tag{23}$$

Provided that a burst is initiated by an alter ($c_1 = A$), the probability that this burst becomes an uninvolved burst of size b is given by the product of $(q_{A|A}F_A)^{b-1}$ and the probability that

an IET is larger than Δt. The latter probability is to guarantee that the burst size is exactly b. Therefore, the numerator on the right-hand side of Eq. (22) is calculated as

$$\sum_{b=1}^{\infty} (q_{A|A} F_A)^{b-1}(1-F) = \frac{1-F}{1-q_{A|A}F_A}. \qquad (24)$$

The denominator on the right-hand side of Eq. (22) can be calculated by enumerating all possible combinations of E and A following the first edit in the burst made by the alter. Let us rewrite each term in the summation over b as $k_b(1-F)$, i.e.,

$$k_b \equiv \frac{\sum_{c_2 \in \{E,A\}} \cdots \sum_{c_b \in \{E,A\}} \Pr[c_2, \ldots, c_b | c_1 = A]}{1-F}. \qquad (25)$$

One gets the following results,

$$k_1 = \Pr[\emptyset|A] = 1, \qquad (26)$$

$$k_2 = \Pr[A|A] + \Pr[E|A] = q_{A|A}F_A + q_{E|A}F_E, \qquad (27)$$

and for $b \geqslant 3$,

$$k_b = q_{A|A}F_A k_{b-1} + q_{E|A}q_{A|E}F_E F_A \sum_{l=0}^{b-3} (q_{E|E}F_E)^l k_{b-l-2}$$
$$+ q_{E|A}q_{E|E}^{b-2}F_E^{b-1}. \qquad (28)$$

In Eq. (28) the first term on the right-hand side contains all cases with $c_2 = A$, while the final term accounts for the case with $c_2 = \cdots = c_b = E$. The second term includes all other cases. Summing up k_b over all b's, we get the denominator on the right-hand side of Eq. (22) as

$$\sum_{b=1}^{\infty} k_b(1-F) = \frac{(1-q_{E|E}F_E+q_{E|A}F_E)(1-F)}{(1-q_{A|A}F_A)(1-q_{E|E}F_E)-q_{E|A}q_{A|E}F_E F_A}. \qquad (29)$$

Then by using Eqs. (12), (19), (21), (24), and (29), for a given timescale Δt we finally obtain the analytic result of the DoI for the entire range of q_E and $q_{E|E}$ as well as for arbitrary functional forms of $P_E(\tau)$ and $P_A(\tau)$:

$$r = \frac{q_E(1-F_E)(1-q_{A|A}F_A)(1-q_{E|E}F_E+q_{E|A}F_E)}{(1-F)q_{E|A}F_E(1-q_{A|A}F_A+q_{A|E}F_A)+q_E(1-F_E)[(1-q_{A|A}F_A)(1-q_{E|E}F_E)-q_{E|A}q_{A|E}F_E F_A]}. \qquad (30)$$

We demonstrate our analytic result in Eq. (30) by adopting the power-law IET distributions with the power-law exponent α, lower bound $\tau_{\min}$, and exponential cutoff τ_c:

$$P(\tau) = \frac{\tau_c^{\alpha-1}}{\Gamma(1-\alpha, \tau_{\min}/\tau_c)} \tau^{-\alpha} e^{-\tau/\tau_c} \theta(\tau - \tau_{\min}), \qquad (31)$$

where $\Gamma(\cdot, \cdot)$ is the upper incomplete Gamma function and $\theta(\cdot)$ is the Heaviside step function. This choice is based on empirical results in the literature [2]. For denoting the parameters for the IET distribution of the ego (the alter) we add the subscript E (A) to those parameters, such as α_E (α_A). The analytic results of the DoI curve for various combinations of parameter values of α_E, α_A, and $q_{E|E}$ are shown as solid curves in Figs. 4(a)–4(c), where in all cases we use $q_E = 0.1$, $\tau_{\min} = 2^2$, and $\tau_c = 2^{10}$ for the ego, and $\tau_{\min} = 1$ and $\tau_c = 2^{15}$ for the alter. These analytic results are successfully confirmed by the simulation results with 100 generated edit sequences of up to $n = 2^{23}$ edits, as depicted by symbols in Figs. 4(a)–4(c).

Next, we numerically calculate the AUC value from the analytic result of the DoI curve in Eq. (30), for which we identify the values of Δt_1 and Δt_2. Since Δt_1 is the maximum value of Δt satisfying $m_{\text{inv}} = 0$, i.e., $f = 1$ in Eq. (20), Δt_1 turns out to be the same as $\tau_{\min}$ for the ego's IET distribution. Δt_2 is determined as the minimum value of Δt satisfying $m_{\text{init}} = 0$ or $p_{\text{init}} = 0$, i.e., $F_E = 1$ by Eq. (19). However, since $F_E(\Delta t) = 1$ can be achieved only when $\Delta t \to \infty$, we instead obtain Δt_2 satisfying the condition $F_E(\Delta t_2) = 1 - 10^{-6}$, which is essentially of the order of τ_c for $P_E(\tau)$.

Figures 4(d)–4(f) shows how the AUC value depends on the parameter values for temporal and editorial correlations. We observe in Fig. 4(d) that the smaller α_E and/or larger α_A lead to the larger AUC values. To compare this finding with the empirical observation in Sec. II C, one needs to understand the relation between the shape of the IET distribution and the burstiness parameter derived from it. Once the functional form of the IET distribution $P(\tau)$ is given as in Eq. (31), the standard deviation σ and the mean μ of the IET are calculated to obtain the value of the burstiness parameter [49]:

$$B \equiv \frac{\sigma - \mu}{\sigma + \mu}, \qquad (32)$$

which can also be derived from Eq. (3) in the case with $n \to \infty$. As a result, we observe the nonmonotonic dependence of B on the power-law exponent α for fixed values of $\tau_{\min}$ and τ_c, as depicted in Fig. 5(a). This implies that a heavier tail of the IET distribution with a smaller value of α does not necessarily lead to the larger value of B, in particular, in the presence of the exponential cutoff to the power-law tail. For the IET distributions with $\tau_{\min} = 1$ and various values of τ_c, we find an increasing behavior of B as a function of α for the range of $\alpha \lesssim 1.5$. Assuming that the empirical IET distributions of the ego and the alters follow the functional form in Eq. (31), we roughly estimate values of the power-law exponent of the empirical IET distribution for the ego and for the alters to find them within the range of $\alpha \lesssim 1.5$, as shown in Figs. 5(b) and 5(c). Conclusively, the analytic result that the smaller α_E and larger α_A lead to the larger AUC value is consistent with the empirical finding of the negative correlation between B_E and AUC values and of the positive correlation between B_A and AUC values.

In Fig. 4(e) we observe that the larger $q_{E|E}$ for a fixed q_E leads to the larger AUC values, which indeed confirms the empirical tendency that the larger AUC values are observed for the stronger editorial correlations as shown in Fig. 2(e). Finally, Fig. 4(f) shows how the AUC value varies when the effects of α_E and $q_{E|E}$ interplay with each other.
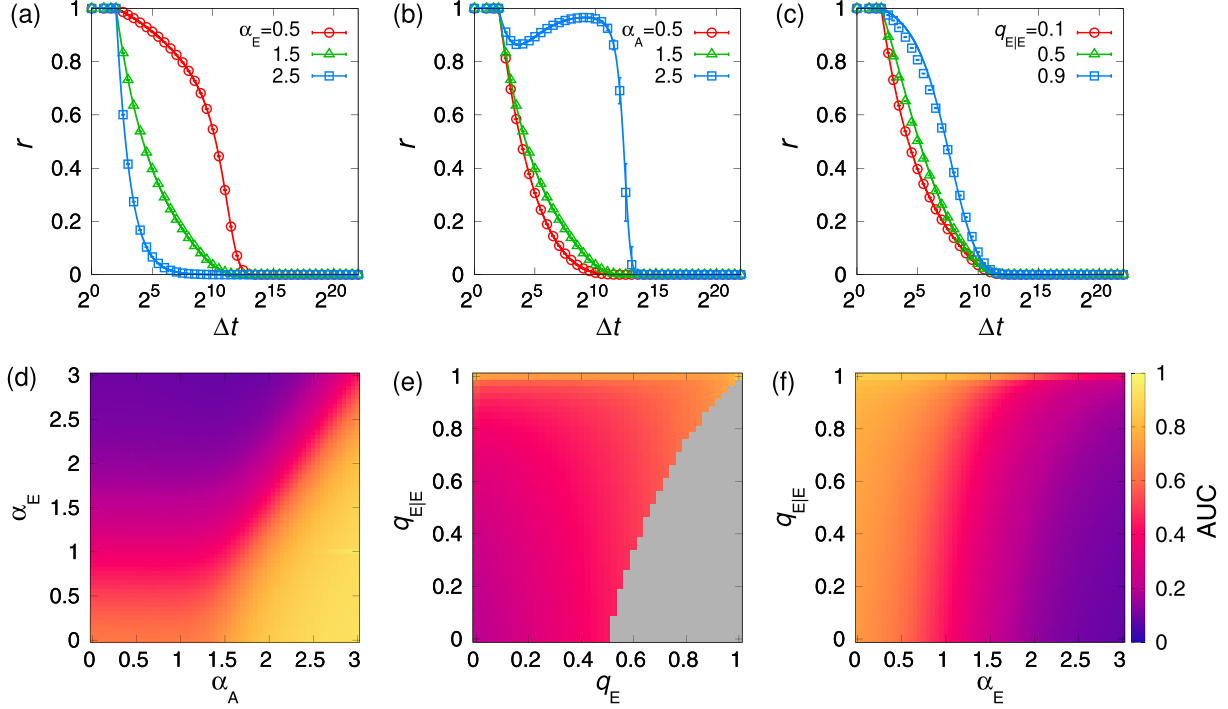
FIG. 4. (a)–(c) Analytic results of the DoI curve of the model in Eq. (30) using $P_E(\tau)$ and $P_A(\tau)$ both in a form of the power-law distribution with exponential cutoff in Eq. (31) (solid curves). In all cases we use $q_E = 0.1$, $\tau_{min} = 2^2$, and $\tau_c = 2^{10}$ for $P_E(\tau)$, and $\tau_{min} = 1$ and $\tau_c = 2^{15}$ for $P_A(\tau)$. We plot the DoI curves for several values of α_E when $q_{E|E} = 0.1$ and $\alpha_A = 1.5$ (a), for several values of α_A when $q_{E|E} = 0.1$ and $\alpha_E = 1.5$ (b), and for several values of $q_{E|E}$ when $\alpha_E = \alpha_A = 1.5$ (c). These analytic results are confirmed by the simulation results from 100 generated edit sequences of up to $n = 2^{23}$ edits (symbols). The error bars denote the standard deviations. (d)–(f) Numerical results of the AUC value calculated from the analytic result of the DoI curve in Eq. (30) for various combinations of the parameter values, i.e., in the space of (α_A, α_E) when $q_E = q_{E|E} = 0.1$ (d), in the space of $(q_E, q_{E|E})$ when $\alpha_E = \alpha_A = 1.5$ (e), and in the space of $(\alpha_E, q_{E|E})$ when $q_E = 0.1$ and $\alpha_A = 1.5$ (f). In all cases we use $\tau_{min} = 2^2$ and $\tau_c = 2^{10}$ for $P_E(\tau)$ and $\tau_{min} = 1$ and $\tau_c = 2^{15}$ for $P_A(\tau)$. The gray area in panel (e) shows the nonexistent parameter space due to conditions in Eqs. (5)–(7).

By devising and analyzing the model we could understand the underlying mechanisms behind the empirical results more rigorously in terms of the effects of temporal and editorial correlations on the DoI curves and their corresponding AUC values. Thus our modeling approach helps us to better understand the issue on individual-driven burstiness versus interaction-driven burstiness in human dynamics.
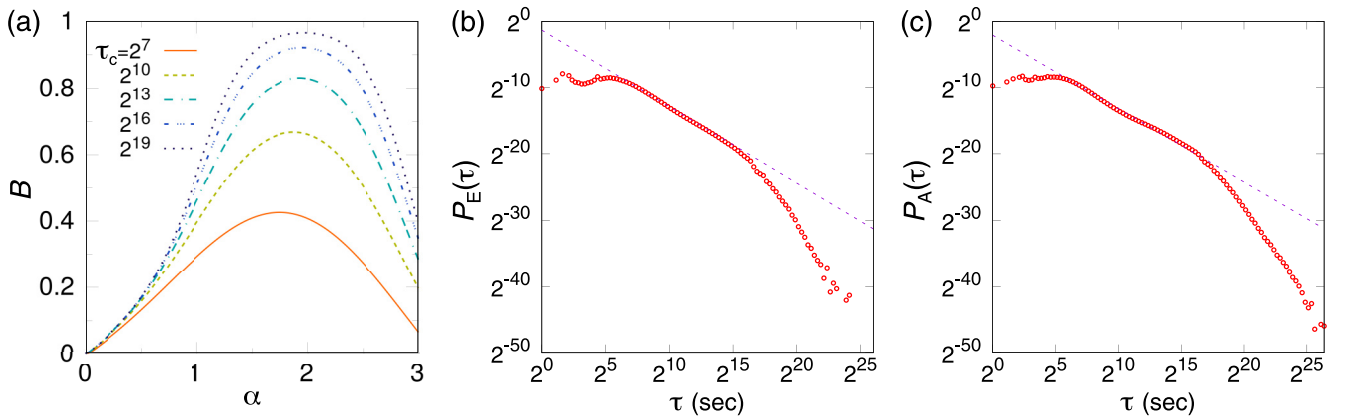

FIG. 5. (a) Nonmonotonic dependence of the burstiness parameter B in Eq. (32) on the power-law exponent α for the IET distribution given by Eq. (31) with $\tau_{min} = 1$ and various values of τ_c. (b) The aggregate distribution of IETs followed by the ego's edits over all article-ego pairs (red circles), with the estimated values of power-law exponent $\alpha_E = 1.15(1)$ (depicted by a dashed line) and burstiness measure $B_E \approx 0.69$. (c) The aggregate distribution of IETs followed by the alters' edits over all article-ego pairs (red circles), with the estimated values of power-law exponent $\alpha_A = 1.11(2)$ (depicted by a dashed line) and burstiness measure $B_A \approx 0.74$.

IV. CONCLUSION

Although the origin of bursty temporal patterns observed in various human behaviors has been extensively investigated, its understanding still remains incomplete [2]. Considering the fact that humans are social beings, one can ask the fundamental question of whether the bursty human dynamics is dominated by the characteristics of individuals or by the interaction between them. In this paper we have addressed this question by analyzing the Wikipedia edit history to see how spontaneous individual editors are in initiating bursty periods of editing and to what extent such editors' behaviors are driven by interaction with other editors in those periods. This question is referred to as individual-driven burstiness versus interaction-driven burstiness.

After detecting the bursty periods or bursts [42] from the edit sequence of each Wikipedia article, we quantify the degree of initiative (DoI) of an individual editor or ego using the statistics of the bursts containing the ego's edits. All other editors are called alters. Since the bursts are detected for a given timescale, the DoI is also a function of the timescale. Then scanning the entire range of timescale in the article, we obtain the DoI curve, from which the area under the curve (AUC) is calculated. The large value of AUC for an article-ego pair implies the dominance of individual-driven burstiness over interaction-driven burstiness and vice versa. By correlating the AUC value with several measures for temporal and editorial correlations, we find the tendency of the AUC values to be larger for weaker (stronger) temporal correlations of the ego (the alters) and/or stronger editorial correlations in the edit sequences. We also successfully confirm these empirical findings by devising and analyzing a model capturing the essential features of edit sequences. Our approach enables us to better understand the origin and underlying mechanisms of bursts in human social dynamics. We also remark that our approach can be applied to any other time series that can be characterized by a sequence of events with both temporal and contextual information. Here the context of the event indicates a situation in which the event occurs, or any other attributes of the event [50,51], such as editors of editing events

in our case and discussion topics of posted messages in online forums [18].

Finally, we discuss possible future works. For the comprehensive description of the temporal and editorial correlations, one can consider higher-order temporal and editorial correlations than those used in our work, e.g., by means of a burst-tree decomposition method for characterizing arbitrary-order temporal correlations in event sequences [27] and ϵ-machine for detecting patterns in the sequence of symbols (e.g., editors in our case) [52,53]. In our work we have focused only on whether the editor of interest initiated bursts or got involved in bursts initiated by other editors. In addition to such initiative behaviors the organization of edits in each burst may also be important to understand the interaction dynamics within bursts, e.g., as studied in Refs. [8,54]. In addition, considering that an editor can edit multiple articles, the same editor may show different initiative behaviors when editing different articles, which can be studied in the future. Further, beyond the article-editor pairs one can extend our approach to encompass the entire web of articles and editors in terms of bipartite temporal networks [29,41,55,56]. The bipartite temporal network for Wikipedia edit history can consist of editing events, denoted by a tuple (i, u, t), indicating that an editor i makes an edit to an article u at time t. Finally, one can also consider information on the interaction between editors from other types of Wikipedia pages, e.g., talk pages [37,38,57], or by means of development tools such as Editor Interaction Analyser [58]. One can gain a much deeper insight into the issue on individual-driven burstiness versus interaction-driven burstiness in the context of collective dynamics; thus one can better understand the origin and underlying mechanisms of bursts in human social dynamics.

[1] A.-L. Barabási, The origin of bursts and heavy tails in human dynamics, Nature (London) **435**, 207 (2005).

[2] M. Karsai, H.-H. Jo, and K. Kaski, *Bursty Human Dynamics* (Springer International Publishing, Cham, 2018).

[3] R. D. Malmgren, D. B. Stouffer, A. E. Motter, and L. A. N. Amaral, A Poissonian explanation for heavy tails in e-mail communication, Proc. Natl. Acad. Sci. USA **105**, 18153 (2008).

[4] H.-H. Jo, M. Karsai, J. Kertész, and K. Kaski, Circadian pattern and burstiness in mobile phone communication, New J. Phys. **14**, 013055 (2012).

[5] S. P. Borgatti, A. Mehra, D. J. Brass, and G. Labianca, Network analysis in the social sciences, Science **323**, 892 (2009).

[6] A.-L. Barabási and M. Pósfai, *Network Science* (Cambridge University Press, Cambridge, 2016).

[7] F. Menczer, S. Fortunato, and C. A. Davis, *A First Course in Network Science* (Cambridge University Press, Cambridge, 2020).

[8] M. Karsai, K. Kaski, and J. Kertész, Correlated dynamics in egocentric communication networks, PLoS One **7**, e40612 (2012).

[9] M. Karsai, M. Kivelä, R. K. Pan, K. Kaski, J. Kertész, A.-L. Barabási, and J. Saramäki, Small but slow world: How network topology and burstiness slow down spreading, Phys. Rev. E **83**, 025102(R) (2011).

[10] G. Miritello, E. Moro, and R. Lara, Dynamical strength of social ties in information spreading, Phys. Rev. E **83**, 045102(R) (2011).

[11] J. Saramäki and E. Moro, From seconds to months: An overview of multi-scale dynamics of mobile telephone calls, Eur. Phys. J. B **88**, 164 (2015).

[12] P. Holme and J. Saramäki, Temporal networks, Phys. Rep. **519**, 97 (2012).

[13] N. Masuda and R. Lambiotte, *A Guide to Temporal Networks*, Series on Complexity Science (World Scientific, New Jersey, 2016).

[14] P. Holme and J. Saramäki, editors, *Temporal Network Theory*, Computational Social Sciences (Springer International Publishing, Cham, 2019).

[15] T. Hiraoka, N. Masuda, A. Li, and H.-H. Jo, Modeling temporal networks with bursty activity patterns of nodes and links, Phys. Rev. Research **2**, 023073 (2020).

[16] J. Yun, S. H. Lee, and H. Jeong, Early onset of structural inequality in the formation of collaborative knowledge in all Wikimedia projects, Nat. Human Behaviour **3**, 155 (2019).

[17] T. Yasseri, R. Sumi, A. Rung, A. Kornai, and J. Kertész, Dynamics of conflicts in Wikipedia, PLoS One **7**, e38869 (2012).

[18] P. Panzarasa and M. Bonaventura, Emergence of long-range correlations and bursty activity patterns in online communication, Phys. Rev. E **92**, 062821 (2015).

[19] G. J. Ross and T. Jones, Understanding the heavy-tailed dynamics in human behavior, Phys. Rev. E **91**, 062809 (2015).

[20] F. Radicchi, Human activity in the web, Phys. Rev. E **80**, 026118 (2009).

[21] J. Yun, S. H. Lee, and H. Jeong, Intellectual interchanges in the history of the massive online open-editing encyclopedia, Wikipedia, Phys. Rev. E **93**, 012307 (2016).

[22] C. Wang, X. Guan, T. Qin, and T. Yang, Modeling heterogeneous and correlated human dynamics of online activities with double Pareto distributions, Inf. Sci. **330**, 186 (2016).

[23] O. Kwon, W.-S. Son, and W.-S. Jung, The double power law in human collaboration behavior: The case of Wikipedia, Physica A (Amsterdam) **461**, 85 (2016).

[24] Y. Zha, T. Zhou, and C. Zhou, Unfolding large-scale online collaborative human dynamics, Proc. Natl. Acad. Sci. USA **113**, 14627 (2016).

[25] Y. Gandica, J. Carvalho, F. Sampaio dos Aidos, R. Lambiotte, and T. Carletti, Stationarity of the inter-event power-law distributions, PLoS One **12**, e0174509 (2017).

[26] J. Jeong and S. Moon, Interval signature: Persistence and distinctiveness of inter-event time distributions in online human behavior, in *Proceedings of the 26th International Conference on World Wide Web Companion: WWW '17 Companion* (ACM Press, Perth, Australia, 2017), pp. 1585–1593.

[27] H.-H. Jo, T. Hiraoka, and M. Kivelä, Burst-tree decomposition of time series reveals the structure of temporal correlations, Sci. Rep. **10**, 12202 (2020).

[28] S. Maniu, B. Cautis, and T. Abdessalem, Building a signed network from interactions in Wikipedia, in *Databases and Social Networks: DBSocial '11* (ACM Press, Athens, Greece, 2011), pp. 19–24.

[29] D. Jurgens and T.-C. Lu, Temporal motifs reveal the dynamics of editor interactions in Wikipedia, in *ICWSM 2012: Proceedings of the 6th International AAAI Conference on Weblogs and Social Media* (AAAI, Palo Alto, CA, 2012), pp. 162–169.

[30] J. Lerner and A. Lomi, The third man: Hierarchy formation in Wikipedia, Appl. Network Sci. **2**, 24 (2017).

[31] J. Lerner and A. Lomi, The free encyclopedia that anyone can dispute: An analysis of the micro-structural dynamics of positive and negative relations in the production of contentious Wikipedia articles, Social Networks **60**, 11 (2020).

[32] T. Iba, K. Nemoto, B. Peters, and P. A. Gloor, Analyzing the creative editing behavior of Wikipedia editors, Procedia: Social and Behavioral Sciences **2**, 6441 (2010).

[33] B. Keegan, D. Gergle, and N. Contractor, Staying in the loop: Structure and dynamics of Wikipedia's breaking news collaborations, in *Proceedings of the Eighth Annual International Symposium on Wikis and Open Collaboration: WikiSym '12* (ACM Press, Linz, Austria, 2012), pp. 1–10.

[34] G. Wu and P. Cunningham, Integration of multiple network views in Wikipedia, Knowledge and Information Systems **45**, 473 (2015).

[35] J. Ashford, L. Turner, R. Whitaker, A. Preece, D. Felmlee, and D. Towsley, Understanding the signature of controversial Wikipedia articles through motifs in editor revision networks, in *Companion Proceedings of the 2019 World Wide Web Conference* (ACM, San Francisco, 2019), pp. 1180–1187.

[36] D. Laniado and R. Tasso, Co-authorship 2.0: Patterns of collaboration in Wikipedia, in *Proceedings of the 22nd ACM Conference on Hypertext and Hypermedia: HT '11* (ACM Press, Eindhoven, Netherlands, 2011), pp. 201–210.

[37] P. Turek, A. Wierzbicki, R. Nielek, A. Hupa, and A. Datta, Learning about the quality of teamwork from wikiteams, in *2010 IEEE Second International Conference on Social Computing* (IEEE, Minneapolis, 2010), pp. 17–24.

[38] M. Jankowski-Lorek, S. Jaroszewicz, Ł. Ostrowski, and A. Wierzbicki, Verifying social network models of Wikipedia knowledge community, Inf. Sci. **339**, 158 (2016).

[39] R. Jesus, M. Schwartz, and S. Lehmann, Bipartite networks of Wikipedia's articles and authors: A meso-level approach, in *Proceedings of the 5th International Symposium on Wikis and Open Collaboration: WikiSym '09* (ACM Press, Orlando, Florida, 2009), pp. 1–10.

[40] G. Wu, M. Harrigan, and P. Cunningham, Characterizing Wikipedia pages using edit network motif profiles, in *Proceedings of the 3rd International Workshop on Search and Mining User-Generated Contents: SMUC '11* (ACM Press, Glasgow, Scotland, UK, 2011), pp. 45–52.

[41] B. Keegan, D. Gergle, and N. Contractor, Do editors or articles drive collaboration?: Multilevel statistical network analysis of Wikipedia coauthorship, in *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work: CSCW '12* (ACM Press, Seattle, 2012), pp. 427–436.

[42] M. Karsai, K. Kaski, A.-L. Barabási, and J. Kertész, Universal features of correlated bursty behaviour, Sci. Rep. **2**, 397 (2012).

[43] English Wikipedia, https://dumps.wikimedia.org

[44] J. Choi, Time series with editorial information for article-ego pairs in Wikipedia edit history, https://github.com/JeehyeChoi/Wikipedia

[45] The ego can be either a human editor or a bot [59]. In our work we focus on the behavior of human editors.

[46] E.-K. Kim and H.-H. Jo, Measuring burstiness for finite event sequences, Phys. Rev. E **94**, 032311 (2016).

[47] H.-H. Jo, Modeling correlated bursts by the bursty-get-burstier mechanism, Phys. Rev. E **96**, 062131 (2017).

[48] H.-H. Jo and T. Hiraoka, Bursty time series analysis for temporal networks, in *Temporal Network Theory* (Springer International Publishing, Cham, 2019), pp. 161–179.

[49] K.-I. Goh and A.-L. Barabási, Burstiness and memory in complex systems, Europhys. Lett. **81**, 48002 (2008).

[50] H.-H. Jo, M. Karsai, J. Karikoski, and K. Kaski, Spatiotemporal correlations of handset-based service usages, EPJ Data Science **1**, 10 (2012).

[51] H.-H. Jo, R. K. Pan, J. I. Perotti, and K. Kaski, Contextual analysis framework for bursty dynamics, Phys. Rev. E **87**, 062131 (2013).

[52] C. R. Shalizi and J. P. Crutchfield, Computational mechanics: Pattern and prediction, structure and simplicity, J. Stat. Phys. **104**, 817 (2001).

[53] C. R. Shalizi, K. L. Shalizi, and J. P. Crutchfield, An algorithm for pattern discovery in time series, arXiv:cs/0210025.

[54] Y. Wu, C. Zhou, J. Xiao, J. Kurths, and H. J. Schellnhuber, Evidence for a bimodal distribution in human communication, Proc. Natl. Acad. Sci. USA **107**, 18803 (2010).

[55] A. Zeng, S. Gualdi, M. Medo, and Y.-C. Zhang, Trend prediction in temporal bipartite networks: The case of Movielens, Netflix, and Digg, Adv. Complex Syst. **16**, 1350024 (2013).

[56] G. Bravo-Hermsdorff, V. Felso, E. Ray, L. M. Gunderson, M. E. Helander, J. Maria, and Y. Niv, Gender and collaboration patterns in a temporal scientific authorship network, Appl. Network Sci. **4**, 112 (2019).

[57] F. B. Viegas, M. Wattenberg, J. Kriss, and F. van Ham, Talk before you type: Coordination in Wikipedia, in *2007 40th Annual Hawaii International Conference on System Sciences (HICSS '07)* (IEEE, Waikoloa, HI, 2007), pp. 78–78.

[58] Editor Interaction Analyser, https://sigma.toolforge.org/editorinteract.py

[59] L. N. Zheng, C. M. Albano, N. M. Vora, F. Mai, and J. V. Nickerson, The roles bots play in Wikipedia, Proc. ACM Human-Computer Interaction **3**, 1 (2019).