Self- and Cross-Excitation in Stack Exchange Question & Answer Communities

Tiago Santos¹, Simon Walk², Roman Kern¹,³, Markus Strohmaier⁴, Denis Helic¹
¹Graz University of Technology, ²Detego GmbH, ³Know-Center, ⁴RWTH Aachen University
tsantos@iicm.edu, s.walk@detego.com, rkern@tugraz.at, markus.strohmaier@cssh.rwth-aachen.de, dhelic@tugraz.at

ABSTRACT
In this paper, we quantify the impact of self- and cross-excitation on the temporal development of user activity in Stack Exchange Question & Answer (Q&A) communities. We study differences in user excitation between growing and declining Stack Exchange communities, and between those dedicated to STEM and humanities topics by leveraging Hawkes processes. We find that growing communities exhibit early stage, high cross-excitation by a small core of power users reacting to the community as a whole, and strong long-term self-excitation in general and cross-excitation by casual users in particular, suggesting community openness towards less active users. Further, we observe that communities in the humanities exhibit long-term power user cross-excitation, whereas in STEM communities activity is more evenly distributed towards casual user self-excitation. We validate our findings via permutation tests and quantify the impact of these excitation effects with a range of prediction experiments. Our work enables researchers to quantitatively assess the evolution and activity potential of Q&A communities.

CCS CONCEPTS
• Human-centered computing → Collaborative and social computing; • Mathematics of computing → Stochastic processes.

KEYWORDS
Q&A communities, Excitation effects, Hawkes processes

ACM Reference Format:

1 INTRODUCTION
Why and how some Question & Answer (Q&A) communities gain traction and attract activity from large numbers of users—while others do not—are questions of theoretical and practical relevance [31, 38]. Understanding how users become active in such systems, and how user activity evolves over time, can be considered an important stepping stone towards better modeling and shaping of online Q&A communities. This will allow to devise novel approaches to guide and encourage activity [12] and to support community managers in their community building efforts [2, 20].

User excitation. In this paper, we investigate self-excitation and cross-excitation of users in Q&A communities. Self-excitation reflects how a user’s own past activity shapes her future activity, while cross-excitation reflects how other users’ past activity influences future activity of a given user. Modeling temporal traces of user activity informs the inference of excitation effects and thus provides a first step towards deeper causal analysis of user excitation.

In the present work, we adopt point processes [7–9]—in particular Hawkes processes [14]—to leverage temporal traces of user activity as latent indicators of self- and cross-excitation. We empirically analyze 69 Stack Exchange Q&A instances where we fit a multivariate Hawkes process model1. With that, we are able to analyze self- and cross-excitation of users across: (i) communities with growing and declining activities, (ii) the topics of the conversations, (iii) activity types (e.g., question, answers), and (iv) activity source (e.g., power user, casual users). Subsequently, we characterize self- and cross-excitation as a function of community age. We then validate, with a range of statistical tests, the excitation effects we uncover, and we quantify their relative importance in the evolution of Q&A communities with a prediction experiment. We illustrate various types of excitation and how they generate user activity in Figure 1.

Findings. Our empirical findings emphasize the need for Q&A communities to maintain a steady core of highly cross-excited power users (i.e. very active users) reacting, particularly in a community’s early stages, to the community as a whole. In thriving communities, casual users (less active users) shape each others’ activity levels via cross-excitation. This suggests that growing communities are, in general, facilitating and embracing less active and casual users, thereby offering low barriers of entry. Additionally, we observe late-stage domination of self-excitation over cross-excitation, meaning that self-driven activity becomes a crucial factor in successful communities. This effect may serve as a long-term growth indicator, as this self-excitation dominance is most prominent in growing communities. Finally, we observe differences in user participation across distinct topics: Q&A communities dedicated to topics in the humanities (such as languages) are more driven by cross-excitation of power users, whereas those in STEM-related fields are not.

With our work we make the following contributions. First, we model self- and cross-excitation effects in successful and unsuccessful Q&A communities. Second, we empirically show how self- and cross-excitation manifests in communities defined by different levels of success and different topics. Third, our validation provides a foundation for building further predictive models of user activity in Q&A communities. Finally, we provide and illustrate an approach

¹We make our code available at https://github.com/tfts/Excitation_in_QA.
that allows community managers to quantitatively compare long-term dynamics of their online Q&A communities—in terms of user excitation—to well-established ones.

## 2 HAWKES PROCESSES

A point process can be broadly defined as a collection of points randomly located in some mathematical space. Temporal point processes employ the real line, representing time, as the underlying mathematical space. For the interested reader, the work by Daley and Vere-Jones [8, 9] and by Cox and Isham [7] are comprehensive references on point process theory.

In practice, temporal point processes model the arrival of discrete events over time with the help of a conditional intensity function $\lambda^*$, a stochastic model for the arrival of the next event given the event history. Hawkes processes [14] are a particular class of temporal point processes, which assume a particular functional form for the intensity function. Specifically, the intensity function of Hawkes processes is itself a stochastic process and it explicitly encodes self-excitation, the increase in intensity caused by past events:

$$\lambda^*(t) = \mu + \sum_{t_i < t} \alpha e^{-\beta(t-t_i)},$$  \hspace{1cm} (1)

where $\mu > 0$ is the baseline intensity independent on the event history, and $\alpha, \beta > 0$ establish the dependence on previous events. In particular, each previous event at time $t_i$ increases the intensity by $\alpha$, the self-excitation factor. We choose to let the intensity jumps exponentially decay at the rate $\beta$, which is a commonly used functional form of intensity decay called an exponential kernel.

Equation 1 describes univariate Hawkes processes, as they consider only the effect of past events in future event times of the same event stream. The multivariate generalization of univariate Hawkes processes includes not only self-excitation but also cross-excitation. Cross-excitation is the intensity increase that an event in one event stream implies in another event stream. More formally, let $N(t) = (N^1(t), N^2(t), \ldots, N^M(t))$ be a simple multivariate point process, where each of the $N^n(t)$ is a counting process in the $i$-th dimension. An $M$-variate Hawkes process with an exponential kernel is defined by the following intensity function:

$$\lambda^m(t) = \mu_m + \sum_{n=1}^M \sum_{t_i^n < t} \alpha_{mn} e^{-\beta_{mn}(t-t_i^n)}.$$ \hspace{1cm} (2)

We write $\mu_m$ as the baseline intensity in dimension $m$, $\alpha_{mn}$ as the cross-excitation on dimension $m$ caused by an event in dimension $n$ and the corresponding decay rate as $\beta_{mn}$. In matrix notation, we write $\mu \in \mathbb{R}^M, \alpha \in \mathbb{R}^{M \times M}$ and $\beta \in \mathbb{R}^{M \times M}$. Linniger [23] provides a more detailed treatment of multivariate Hawkes process theory.

Samples from multivariate Hawkes processes, which can be obtained via Ogata’s thinning algorithm [27], generate self-excitation and cross-excitation effects of the kind depicted in Figure 1, where users A and B correspond to two dimensions of a Hawkes process and their event rate to the Hawkes process intensity. We observe, in both dimensions, intensity peaks corresponding to the sampled events, and we note the intensity decays exponentially until converging to the baseline intensity level $\mu$.

### Fitting Hawkes processes

Given an observed sequence of events $\{t_i\}$, we fit the parameters of Hawkes processes by maximizing its log-likelihood. Closed form expressions for the log-likelihood can be derived for many different types of intensity function kernels, including the exponential kernel that we assume. For a given $\beta$, all other parameters of the process may be estimated by maximizing the log-likelihood via well-known convex optimization methods.

---

**Figure 1:** Excitation types. We distinguish between three drivers of user activity excitation: (a) Baseline, a constant base event rate level, (b) Self-Excitation, a proxy for increased propensity by a user to be active in the future following her past activity, and (c) Cross-Excitation, a boost to event rate triggered by activity of other users. The upper row of this Figure depicts the links between users A and B per excitation type, and the two lower rows the corresponding event rate as it reacts to the user's own and others' activity events, marked by the trees below each event rate line and colored by the corresponding user. The first three excitation types cover excitation components which combine to form (d) empirical excitation. This work characterizes and quantifies how each type of excitation manifests itself in Q&A communities and how excitation strength changes over time as a Q&A community develops.
such as Levenberg-Marquardt [21]. However, fitting $\beta$-s is a challenging task, since the likelihood functions of Hawkes processes with exponential kernels are either flat around the optimal $\beta$ (see, for example, Upadhyay et al. [36]) or, in some other formulations of the kernel function, even non-convex in $\beta$. In this paper, we propose an effective Bayesian hyperparameter optimization step, which allows fitting the decay-related and then the remaining parameters of a Hawkes process. Assuming $\beta_{m,n} = \beta, \forall 1 \leq m,n \leq M$ [12, 35], we apply the Tree of Parzen Estimators approach, as described by Bergsma et al. [4], on the convex optimization routine of log-likelihood for a given set of event sequences to estimate $\beta$. We perform 15 runs of the Bergsma et al. algorithm and keep the $\beta$ yielding highest likelihood, since this effectively allows for convergence even in the presence of flat plateaus around local maxima of Hawkes likelihood as a function of $\beta$. Finally, using the learned $\beta$, we fit $\mu$ and $\alpha$.

Furthermore, practical fitting of Hawkes processes requires the fits to be done on stationary [23] periods of the corresponding count time series. Stationarity, in this context, refers to translation-invariance in the Hawkes process distribution, which implies a linear growth in the associated time series of events counts over time. However, in the time series representing activity in the Q&A communities we work with, we observe a range of non-stationary phenomena: exponential growth and decline and other sudden structural changes, such as level jumps. Therefore, we need to restrict the fitting procedure to stationary subsequences of an observed event stream. To that end, we use the time series structural change detection algorithm devised by Zeilis et al. [45]. Given a linear regression model, this algorithm returns optimal points in time for structural change in the regression model’s fit to a given input time series. Using a constant regression model allows us to detect level changes in an input event count time series, and thus to segment it into stationary subsequences.

3 EXPERIMENTAL SETUP

We study user self- and cross excitation by empirically analyzing Stack Exchange instances. We distinguish activity in these datasets by two aspects: (i) activity content, which we define as questions and answers (with the latter including answers and comments), and (ii) activity source, meaning whether it originated from highly active (power) or less active (casual) users. Further, we explore the differences in excitation across communities, which we group according to two criteria: (i) growth pattern, and (ii) topical focus. Datasets. Stack Exchange encompasses several Q&A communities, termed Stack Exchange instances, with each dedicated to Q&A on a single topic, such as computer science, the English language or movies. We extract user activity in all 159 Stack Exchange instances\(^2\) (as of June 2017) as the timestamps of users’ activity events: posts (i.e. questions) and replies (i.e. answers and comments). In a first step, we consider these instances’ complete history, which spans the period from August 2008 to June 2017 and comprises a total of 22 million events. However, our analysis is independent from the calendar date a Stack Exchange instance originated, as we map the inception of each instance to a time scale starting at zero.

Table 1: Dataset characteristics. We show the datasets per group sorted by activity growth (top and bottom three growth percentages for group shown in parenthesis), the total number of datasets per group (#), as well as ranges for a number of descriptive statistics: the activity total as the sum of all questions and answers, the age in years and the total growth as a percentage of the level of the first subsequence found by Zeilis et al.’s algorithm. We observe a clear separation in strongly positive and negative growths (and thus also total activity) in the major distinction we draw between datasets, growing vs. declining. This distinction is remarkably less pronounced in STEM vs. humanities instances, which both feature positive and negative growths.

<table>
<thead>
<tr>
<th>Dataset Group</th>
<th>Datasets</th>
<th>#</th>
<th>Activity total</th>
<th>Age (years)</th>
<th>Growth (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stack Exchange Growing</td>
<td>electronics (757.62%), ru (736.42%), codegolf (510.06%), chemistry, sharepoint, academia, puzzling, tex, codereview, blender, unix, money, go, ux, crypto, security, stats, salesforce, dba, wordpress (182.28%), opendata (174.69%), askubuntu (169.29%)</td>
<td>22</td>
<td>[7987, 1489384]</td>
<td>[3.08, 7.83]</td>
<td>[169.29, 757.62]</td>
</tr>
<tr>
<td>Stack Exchange Declining</td>
<td>boardgames (~28.53%), fitness (~34.56%), sound (~35.01%), productivity, tridon, parenting, pets, craftfans, webapps, spanish, cooking, ham, bricks, gardening, cth0ry, expressionengine, pm, skeptics, sustainability, genealogy (~80.26%), ebooks (~81.52%), stackapps (~82.7%)</td>
<td>22</td>
<td>[3301, 117474]</td>
<td>[3, 7.75]</td>
<td>[−82.7, −28.53]</td>
</tr>
<tr>
<td>Stack Exchange STEM</td>
<td>electronics (757.62%), chemistry (473.48%), stats (199.18%), biology, datascience, physics, astronomy, cs, space, cogsci, earthscience, engineering, reverseengineering (0.00%), softwareengineering (~21.28%), sound (~35.01%)</td>
<td>15</td>
<td>[15759, 745764]</td>
<td>[2.41, 8.75]</td>
<td>[−35.01, 757.61]</td>
</tr>
<tr>
<td>Stack Exchange Humanities</td>
<td>philosophy (122.45%), english (117.76%), chinese (23.17%), music, german, mythology, portuguese, christianity, esperanto, arabic, russian, writers, buddhism (~26.62%), french (~27.91%), spanish (~50.10%)</td>
<td>15</td>
<td>[87, 896631]</td>
<td>[0.17, 6.83]</td>
<td>[−50.10, 127.47]</td>
</tr>
</tbody>
</table>
Before we group Stack Exchange instances according to our two criteria, growth patterns and topical focus, we discard all datasets with less than ten events in any period of three months, which ensures that we have enough events for the fitting procedure.

**Growth pattern.** In our first comparison by growth pattern, we analyze the first three years of existence of Stack Exchange instances, so we exclude datasets with durations shorter than three years. To better distinguish excitation effects driving overall activity increase or decrease in these communities, we then extract, from the remaining datasets, two groups of strongly growing and strongly declining datasets. The extraction criterion stems from our application of Zeilis et al.'s algorithm to find level structural changes in the time series of total activity count per month: We define a dataset as strongly growing (declining), if the percentage change in structural level from the first fitted window to the last fitted window is in the 80th (20th) percentile over all datasets. The grouping of Stack Exchange instances into growing and declining yields two groups of 22 datasets each, of which we provide descriptive statistics in Table 1. Note that the growing (declining) group only includes instances with strongly positive (negative) growth.

We plot the total monthly event counts for a selected dataset from each dataset group in Figure 2 to exemplify their activity curves and the detected structural level changes. Often, there are prolonged periods of stagnancy in one structural level, both in growing as well as declining datasets. Typically, such periods vary in length.

**Topical focus.** For the topical comparison, we study Stack Exchange instances dedicated to STEM (i.e. science, technology, engineering and mathematics) and humanities topics. To that end, we randomly picked a set of 15 Stack Exchange instances we manually classified as STEM topics, and another 15 as humanities. The instances in these two groups vary in size and age, and feature no distinctive growth patterns, although some humanities instances are smaller and have shorter overall durations than STEM instances (cf. Table 1). In this comparison, we also analyze instances’ first three years, but we do not impose a minimum duration, which leads to fewer than 15 instances per group later in time. However, the number of instances per group remains comparable over time and reaches a minimum of nine instances per group by the third year.

**Hawkes process application.** In the Stack Exchange instances, we distinguish between more active and less active users, which we term power users and casual users. This definition mainly distinguishes a core of remarkably engaged power users typically found in Q&A communities [13, 24, 41] from casual users. Thus, for each dataset individually, we count the total activity per month per user, and per activity type (question or answer) and postulate that power users are those in the 90th percentile of most active users for that month. This implies that this monthly group of power users is ever-changing, as users join and leave the communities or as the users’ intrinsic motivation to contribute content rises and falls over time. Note that our results changed only minimally with different percentile thresholds (i.e., 85th and 95th) for power user classification.

To measure self- and cross-excitation per user and activity event type, we map the event stream of question and answer activity to four Hawkes process dimensions: questions by power users, questions by casual users, answers by power users and answers by casual users. For each such dimension, we work with the corresponding event timestamps at the resolution of a second.

We then follow the procedure outlined in Section 2 to fit four-dimensional Hawkes processes to each dataset group (Stack Exchange instances in the groups declining vs. growing and STEM vs. humanities). For each dataset group comparison, we begin by fitting overall $\beta$ for all datasets over the first three years of their existence. Then, we perform structural level change fits on the total monthly event count, and observe a minimal window length of five months. According to our experimentation with different window lengths, specifically two to six months, we find a window length of three months is long enough to ensure we have enough events per window and do not overfit a particular window, while also short enough to capture granular changes in the evolution of the underlying Hawkes process distribution. Hence, we set the constant window length to three months (a quarter).

To measure variability in the evolution of the fitted models, we bootstrap, with 100 repetitions, the fitting procedure of all Stack Exchange instances per dataset group per window. From the resulting bootstrap distribution, we compute 95% confidence intervals for the mean value of each fitted Hawkes process parameter. We display the confidence intervals as error bars in Figures 3, 4 and 5.
4 EXCITATION EFFECTS

4.1 Comparison by growth pattern

The $\beta$ value, fitted over 44 datasets in the growing vs. declining comparison and the whole three year period, is 2.288, corresponding to an intensity half-life of about 0.3 hours (meaning that the intensity jump of magnitude $\alpha$ caused by either self-excitation or cross-excitation decays to $\alpha/2$ after about 18 minutes). With a single constant $\beta$, we restrict our model to capture distributional changes in terms of baseline, self- and cross-excitation intensities, allowing us to focus on these factors as direct proxies for the role of different user groups in overall activity intensity over time.

We visualize the evolution of all baseline parameter values for questions and answers by power users and casual users in growing (orange) and declining (blue) instances in Figure 3. We depict the corresponding self- and cross-excitation parameter values in Figure 4. Note that we employ the same scales throughout both Figures for better comparison.

Low baseline intensities. The Figures 3a through 3d show the baseline intensities ($\mu$) fitted over time. We observe roughly constant baseline intensities throughout the whole period, for both growing and declining instances. Furthermore, the baseline intensities are rather low, especially in comparison with the self-excitation and cross-excitation effects ($\alpha$) depicted in Figure 4.
Finding: Constant and low baseline intensities suggest Q&A communities thrive off self- and cross-excitation, representing interaction between different user (power and casual) and activity types (questions & answers), rather than featuring constant levels of activity over time, independent from other activity dimensions.

Early Power User Cross-Excitation (●). We continue our analysis with the Hawkes process dimension with the highest intensity values: intensity in answer activity by power users (first row in Figure 4, i.e. Figures 4a through 4d). We observe, in the early stages of growing instances, high impact of questions by both power as well as casual users on answer activity by power users (see yellow region highlighted in Figures 4b and 4d). In contrast, in declining instances, especially in the yellow highlighted region, questions by both types of users elicit declining numbers of answers by power users over time, and self-excitation in answers by power users dominates over all temporal windows. Regarding question activity by power users (second row in Figure 4, i.e. Figures 4e through 4h), there is a clear prevalence of self-excitation intensity with respect to other cross-excitation intensities. However, we observe, albeit minor, differences in the short and medium terms between growing and declining instances, as power users in growing instances are more encouraged to participate with new questions as a response to questions by casual users and answers by both (see inlines of Figures 4e, 4g, 4h).

Finding: These observations suggest strong activity by power users, as a response to questions by both power users as well as, crucially, casual users, is related to increased growth in the early stages of Q&A communities. This finding suggests the importance of an active core of users to jumpstart Q&A community development.

Late Casual User Cross-Excitation (●). In Figure 4’s third row (Figures 4i through 4l), we highlight another type of effect: Answers and discussion by casual users is driven strongly by questions also from casual users, especially in the long-term as highlighted by the green region of Figure 4l. The main difference between growing and declining instances in this dimension is, besides the intensity magnitude difference, that this cross-excitation effect loses importance in the long-term in the declining instances, while overall it does not in growing instances. We point out one interesting effect in the fourth row (Figures 4m through 4p), which depicts the question intensity dimensions of casual users: In the long term, there is a small increase in questions by casual users after answers also by casual users in growing Stack Exchange instances (cf. quarters eight through twelve of Figure 4o).

Finding: Long-term cross-excitation from questions on answers by casual users is a key factor present in growing Stack Exchange instances and lacking in declining ones. We find contributions by casual users thus attract more participation by casual users, likely helping to sustain and even enhance activity levels. Hence, we identify openness from the community towards casual users in the form of healthy interaction between them as a sign of enduring community growth.

Late Stage Self-Excitation (●). In the diagonal of Figure 4, consisting of Figures 4a, 4f, 4k and 4p, we observe strong and growing self-excitation effects, which dominate over cross-excitation effects in the long-term. We indicate long-term with the pink region marking quarters 8 to 12, the last five quarters we fit. We note this effect is most predominant in growing communities. Further, for growing instances, notice that, for a given dimension (e.g. answers by power users, Figures 4a through 4d), the timing of the surge in self-excitation coincides in general with a decline in cross-excitation.

Finding: In growing Stack Exchange instances, we attribute the phenomenon of higher long-term self-excitation to steadily growing arrivals of questions and answers from power and casual users. As users react to a constantly and regularly growing pool of questions and answers, this makes distinction of direct interaction between single questions and corresponding answers harder over time. The timing of this self-excitation surge may be of particular interest for Q&A community managers, who may be concerned about growth should they not observe this effect by the community’s third year.

4.2 Comparison by topic

The comparison between STEM and humanities instances of Figure 5 shares a few commonalities with our previous findings on the growing and declining instance comparison: roughly constant and relatively low baseline intensities (not depicted due to limitations in space) and comparatively high long-term self-excitation. In this comparison, we obtained β = 2.067, which corresponds to an intensity half-life of 0.33 hours. These values are comparable to the ones we obtained previously, which may indicate a universal pattern of user activity decline across Stack Exchange instances.

Power User Cross-Excitation (●) vs. Casual User Self-Excitation (●). In the light blue highlighted region of Figure 5d, we stress the notable role of answer activity by power users in humanities instances. We observe answers by power users after questions from casual users is notably higher in humanities instances than in STEM instances (see Figure 5d). With the light orange region of Figure 5k, we underline the counterpart in casual user activity: There are higher long-term intensities in self-excitation of answers by casual users in STEM instances as compared to humanities.

Finding: In comparison with STEM Stack Exchange instances, humanities Stack Exchange instances are more reliant on cross-excitation by power users to address questions by both types of users. We observe more power user centric interactions in Stack Exchange instances in the humanities, while activity in STEM Stack Exchange appears more focused on casual users. Higher long-term self-excitation by casual users in STEM instances indicates stronger interactions between casual users. In turn, casual user activity is less dependent on power users in these instances. Overall, this finding suggests the existence of topic-dependent user type structures, which can be cast as measurable goals for community managers.

5 EVALUATION

In this section, we assess whether differences in excitation effects we observe in the evolution of growing vs. declining (STEM vs. humanities) instances result by chance or if there is some causal link between excitation, as measured with the Hawkes processes, and community growth (topical focus). Moreover, we evaluate the sizes of the observed effects by quantifying their impact on the future user activity.

Comparison of activity distributions. While the comparison of growing vs. declining instances aims to distinguish excitation effects in instances of increasing vs. decreasing and thus different total
activity volumes, the STEM vs. humanities comparison is intended towards providing decoupled effects, which ideally should not be confounded with the excitation effects of growing vs. declining instances. However, in the STEM vs. humanities instance comparison, we highlight long-term self-excitation in answers by casual users, an effect which could be similar to the late stage self-excitation of growing vs. declining instances. Furthermore, if humanities instances simply featured overall higher answer-based activity levels by power users than in STEM instances, power users would also likely react stronger to questions by casual users, as opposed to them being an inherently more important backbone to questions by casual users.

Hence, we verify if the total answer-based activity distributions of both user types are similar in STEM and humanities instances. We compare the sample distributions of answers-based activity by power (and separately casual) users in STEM vs. humanities instances with the Kolmogorov-Smirnov two-sample test for their equality. As this test results in a p-value of 0.3855 (0.2305) for power (casual) users’ activity distributions, we conclude there is not enough evidence to reject the null hypotheses of the probability distributions being equal at all usual significance levels. In turn, this test result indicates that the power and casual users’ activity distributions are comparable, which supports our finding regarding the importance of the role power (casual) users play in humanities (STEM) Stack Exchange instances.

**Permutation tests.** To assess the significance of the excitation effects we conduct the following permutation test. First, we randomly permute the association of event types (questions by power users, questions by casual users, answers by power users and answers by casual users) to the corresponding time stamps per time window.
This procedure keeps the amount of events per event type constant, but destroys the temporal connection between event types. Then, we refit the multivariate Hawkes process over windows of these permuted event streams, and we repeat these two steps 100 times. Finally, we compare the difference in mean Hawkes process parameter values fitted on the permuted event streams to the original ones. If there is a notable difference between them, this indicates that growth (or the evolution of the topological instances) does not come about by chance, but that differences in self-excitation and cross-excitation between growing vs. declining (STEM vs. humanities) communities play an important role in their temporal evolution. We depict the result of these permutation tests in Figure 6, in which previously described differences between growing vs. declining and STEM vs. humanities Stack Exchange instances are all either remarkably weaker or non-existent. We arrive at similar conclusions as previously: answers by power users, questions by power users, and for each three-month time window (quarter) decling the absolute difference of growing vs. declining (STEM vs. humanities) Hawkes process parameter values fitted on the original event streams with the distribution of absolute differences in parameter values obtained on permuted event streams. If the difference in original values is extreme in relation to the distribution of permuted values in the effects’ time spans, then this is further evidence the effects we observe are unlikely to arise by chance. We quantify “extreme” with the p-value, in this case the proportion of values from the permuted difference distribution greater than the original difference.

**Prediction experiment.** To quantify the impact of the observed excitation effects on future activity, we design the following prediction experiment. For each three-month time window (quarter) and for each growing and declining Stack Exchange instance, we fit three variants of the Hawkes process, with the same four dimensions as previously: answers by power users, questions by power users, and for each three-month time window (quarter) decling the absolute difference of growing vs. declining (STEM vs. humanities) Hawkes process parameter values fitted on the original event streams with the distribution of absolute differences in parameter values obtained on permuted event streams. If the difference in original values is extreme in relation to the distribution of permuted values in the effects’ time spans, then this is further evidence the effects we observe are unlikely to arise by chance. We quantify “extreme” with the p-value, in this case the proportion of values from the permuted difference distribution greater than the original difference.

### Table 2: Kolmogorov-Smirnov (K-S) distance between predicted and real interevent times, by predicted quarter, effect type, Hawkes process dimension and model variant. Lower K-S values are better. Distance values marked with an asterisk correspond to not rejecting equality of distributions as previously: answers by power users, questions by power users, and for each three-month time window (quarter) decling the absolute difference of growing vs. declining (STEM vs. humanities) Hawkes process parameter values fitted on the original event streams.

<table>
<thead>
<tr>
<th>Prediction Quarter</th>
<th>Early Power User Cross-Excitation</th>
<th>Late Casual User Cross-Excitation</th>
<th>Late Stage Self-Excitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Answers by Power</td>
<td>0.26 0.26 0.29 0.27 0.24</td>
<td>0.24 0.25 0.23 0.24</td>
<td>0.24 0.25 0.23 0.24</td>
</tr>
<tr>
<td>Questions by Power</td>
<td>0.3 0.28 0.25 0.25 0.24</td>
<td>0.23 0.24 0.23 0.24</td>
<td>0.23 0.24 0.23 0.24</td>
</tr>
<tr>
<td>Answers by Casual</td>
<td>0.24 0.22 0.25 0.21 0.21</td>
<td>0.2 0.18 0.17 0.18</td>
<td>0.2 0.18 0.17 0.18</td>
</tr>
<tr>
<td>Questions by Casual</td>
<td>0.23 0.16 0.18 0.13 0.12</td>
<td>0.12 0.14 0.13 0.12</td>
<td>0.12 0.14 0.13 0.12</td>
</tr>
<tr>
<td>Excitation Effects Removed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Answers by Power</td>
<td>0.15 0.13 0.14 0.13 0.13</td>
<td>0.1 0.11 0.11 0.11 0.12</td>
<td>0.43 0.42 0.41 0.42</td>
</tr>
<tr>
<td>Questions by Power</td>
<td>0.28 0.27 0.27 0.26 0.25</td>
<td>0.22 0.24 0.23 0.25</td>
<td>0.3 0.33 0.3 0.31</td>
</tr>
<tr>
<td>Answers by Casual</td>
<td>0.22 0.12 0.16 0.15 0.16</td>
<td>0.15 0.14 0.13 0.14</td>
<td>0.42 0.4 0.39 0.4</td>
</tr>
<tr>
<td>Questions by Casual</td>
<td>0.24 0.19 0.21 0.17 0.17</td>
<td>0.15 0.16 0.16 0.16</td>
<td>0.19 0.19 0.19 0.19</td>
</tr>
<tr>
<td>Full</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Answers by Power</td>
<td>0.16 0.12 0.12 0.11 0.11</td>
<td>0.1 0.11 0.11 0.11 0.11</td>
<td>0.1 0.11 0.11 0.11 0.11</td>
</tr>
<tr>
<td>Questions by Power</td>
<td>0.28 0.24 0.23 0.22 0.22</td>
<td>0.22 0.24 0.22 0.24</td>
<td>0.22 0.24 0.22 0.24</td>
</tr>
<tr>
<td>Answers by Casual</td>
<td>0.23 0.11 0.15 0.14 0.15</td>
<td>0.12 0.11 0.11 0.11 0.11</td>
<td>0.12 0.11 0.11 0.11 0.11</td>
</tr>
<tr>
<td>Questions by Casual</td>
<td>0.25 0.16 0.18 0.13 0.13</td>
<td>0.12 0.13 0.13 0.12</td>
<td>0.12 0.13 0.13 0.12</td>
</tr>
</tbody>
</table>
To further validate these results we perform another prediction experiment with a fourth variant of Hawkes processes. In this variant, we fit self-excitation only models by setting all cross-excitation parameters to zero. These additional experiments with a self-excitation model confirm previous observations: A model with only self-excitation achieves performances (as measured by the K-S distance and the event count RMSE) in general on par with those of the Full model and, in the Late Stage Self-Excitation effect, even surpassing its performance slightly.

For all model variants we come to comparable conclusions when measuring the RMSE between simulated and real event counts. Limited by space, we summarize these results: The average RMSE of the Full model is 638.17 events, an improvement of 59.91% (43.09%) upon the Excitation Effects Removed (Baseline) model.

6 LIMITATIONS

Although we experimented with slightly different percentiles in the user type distinction and the instance characterization and obtained qualitatively similar results, we recognize those are arbitrary thresholds, which impact the results if changed significantly. Our results are more robust to changes in window size of the activity event stream of the Q&A communities (e.g. to two, four or five months), since this hyperparameter controls for the granularity of our results. Nevertheless, the Hawkes process model itself could include time-varying parameters, as an alternative to this repeated fitting procedure we apply over fixed time windows.

The effect time spans we propose, namely a one-and-a-half-year-long early stage and a late stage starting in the last quarter of the second year of a Q&A community, stem from our empirical observations of large differences in excitation in specific temporal segments in the community comparisons. Pinpointing exact transition dates is beyond the scope of this work, as we focus on learning temporal user excitation effects.

We acknowledge that mapping each user’s questions and answers event streams to a Hawkes process dimension may be a more realistic model. However, we argue that such a model would suffer from sparsity, high dimensionality and higher computational cost. Further, such a model might also not improve the excitation effect characterization, as it would also struggle with distinguishing sources of self- and cross-excitation in high-activity regimes.

Note that we avoid a discussion of how casual users become power users with our characterization of power users as the most active each month, regardless of their histories. We believe engagement reward systems such as badges play an important role in casual user’s development in particular and user excitation in general [20], but we leave a detailed investigation of the role of reward systems on excitation effects for future work.

We also caution that our work indicates a temporal link between (i) specific community structures in terms of user types and their excitation and (ii) the overall development of activity volume in a Stack Exchange community. This work does not establish causality.

7 RELATED WORK

Research on Q&A communities. There is a considerable amount of authors [1, 10, 13, 24, 33, 41] analyzing the roles different types of users play in Web communities such as Q&A websites. In addition,
several authors surveyed the motivation and behaviour of individual users [17, 26] of Q&A communities. While Mamykina et al. [24] and Furtado et al. [13] concentrate on uncovering and studying the roles of user types present in thriving Q&A communities, Danescu-Niculescu-Mizil et al. [10] and Yang et al. [41] explicitly focus on specific user types in Web communities and the user types’ static and temporal characteristics. More broadly, Yang et al.’s work is part of a larger body of literature [28, 30, 46] on identifying experts in Q&A websites and characterizing their behavior. Our work leverages a comparable user type characterization to infer properties about the temporal evolution of communities themselves.

In approaches methodologically related to Zhang et al.’s [46], multiple authors [3, 25, 31, 38, 39, 43] study evolution dynamics of Web communities by relying on an explicit description of networks underlying a given Web community, and these networks often serve as a basis for dynamical systems models of the communities. In their study of Quora, another Q&A website, Wang et al. [39] analyze the role different social network structures play in Quora’s community growth. Ribeiro [31] and Walk et al. [38] model users and activity in diverse Web communities including Q&A communities, with the former focusing on growth and decline of communities and the latter on the model’s implications for self-sustainability in a community’s activity. Matsubara et al. [25] and Zang et al. [43] study information diffusion and growth dynamics of Web communities.

Similarly to Matsubara et al.’s, Walk et al.’s and Zang et al.’s work, in this paper we also model growth and interaction dynamics of Q&A communities, but we do not assume an underlying network. We focus rather on excitation between groups of users, which we distinguish not on their expertise but on their overall activity levels. Furthermore, by encoding community lifecycles in Hawkes processes fitted to sequences of time windows, we extend the empirical discussion of Web community lifecycles [15, 42] and the critical mass literature [29, 32] to the Q&A community domain with measurable results.

**Applications of Hawkes processes.** Hawkes processes and their variations, as models for event streams with unequally spaced events in time, have found wide application in literature on different aspects of Web phenomena [12, 16, 36, 37, 44, 47, 48]. One such topic regards content popularity dynamics, in particular how to predict the influence of internal and external aspects of activity in social networks [12] and reshare popularity of items on the Web [47]. To infer causal links between users and user influence from user activity in social networks, Ver Steeg and Galstyan [37], Iwata et al. [16] and Zhou et al. [48] propose point process-related approaches, which cope with high dimensionality in number of users. Further, Upadhyay et al. [36] model the crowdlearning process of Stack Overflow users and characterize different user types by their expertise and learning curves. The work by Zang et al. [44] models and predicts the growth dynamics of individuals’ ties in social networks and predicts its evolution.

Our work draws inspiration and methodological know-how from all above mentioned papers to expand on a topic closely related to Zang et al.’s: the development of not just the relatively small circle of an individual’s social ties, but of excitation and interaction of user groups in Q&A communities. Furthermore, we contribute, to the growing body of work on fitting Hawkes process kernels [5, 22, 40, 49], a parsimonious Bayesian hyperparameter optimization method for fitting the decay parameter of exponential kernels in Hawkes processes. Finally, our extension of this fitting method to non-stationary multivariate event streams enables the extraction of temporal excitation effects from Q&A communities.

8 CONCLUSIONS

**Summary.** In this work, we modeled self- and cross-excitation in Q&A communities along several dimensions, including activity type, user engagement level, growth path and topical focus of a given community. We approached this task by fitting multivariate Hawkes processes to stationary temporal segments of Q&A communities’ activity volumes. We found stronger cross-excitation of power (casual) users in early (late) stages of growing communities when compared to communities with declining total activity. Further, in growing communities, we observed self-excitation dominates in the long-term. Moreover, we uncovered strong long-term cross-excitation by power (casual) users in Q&A communities dedicated to topics in the fields of the humanities (STEM). We validated the presence of these excitation effects with statistical and permutation tests and we quantified their strength via prediction tasks.

**Implications.** Our work can support Q&A community managers in their ambition to promote sustainable community structures. To jumpstart community growth in its first six months, engaging a core of power users, for example in community building efforts, appears to be of crucial importance. In the medium- to long-term, we find community developers should carefully monitor and foster participation rather than casual users. While literature on critical mass in Web communities [29, 32] and studies on the user mix in Wikipedia [18, 34] also support this recommendation, we can afford further advice, as our casual cross-excitation analysis specifically underlines the importance of interaction between casual users. In practice, we believe adjusting reward or badge systems to encourage contributions by casual users, perhaps by welcoming newcomers or by easing their adjustment to community rules, would be of value to community development. Furthermore, community managers, which have not observed a surge in self-excitation by the third year of their communities, may have reason to concern over growth. Such excitation effects should be carefully monitored, as Q&A community growth may come at the cost of other community parameters [11]. Furthermore, our results indicate concrete implementations of these suggestions should depend on community topic, as it impacts excitation effects. Overall, our findings thus highlight the impact of timing in the user mix development.

**Future work.** Comparing other Q&A communities would allow to further generalize the results we obtained on Stack Exchange communities. Our work can be extended to uncover excitation effects in other domains, such as of Q&A instances in other languages or of other contribution types (e.g. open-ended vs. focused question), as our proposed approach is generic and can be readily extended.

**ACKNOWLEDGMENTS**

We thank the anonymous reviewers for their valuable feedback on the manuscript. Tiago Santos is a recipient of a DOC Fellowship of the Austrian Academy of Sciences at the Institute of Interactive Systems and Data Science of the Graz University of Technology.