The Bayesian Echo Chamber: Modeling Influence in Conversations

Fangjian Guo
Duke University
Durham, NC, USA
guo@cs.duke.edu

Charles Blundell*
Gatsby Computational Neuroscience Unit
University College London
London, UK
blundellc@gmail.com

Hanna Wallach†
Microsoft Research
New York City, NY, USA
hanna@dirichlet.net

Katherine Heller
Duke University
Durham, NC, USA
kheller@stat.duke.edu

Abstract

We present the Bayesian Echo Chamber—a new Bayesian model for social interaction data that captures both temporal dynamics and linguistic content. This model is designed to discover latent influence relationships between individuals. We explore two forms of influence by looking at two types of social behavior in group discussions: conversational turn-taking, and linguistic accommodation. These notions of influence are captured via a temporal component and a linguistic component, which use a multivariate Hawkes process and a discrete analog, respectively. The model permits a fully Bayesian inference algorithm. We showcase our model's ability to discover latent influence relationships using transcripts of natural conversations, arguments heard by the US Supreme Court, and the 1957 movie "12 Angry Men." We provide an analysis of the inferred relationships.

1 Introduction

As increasing amounts of data detailing human social interactions become available, largely through online sources, researchers strive to find new ways of leveraging these data in order to learn more about human social behavior. Most social processes, in which individuals or groups of individuals interact with each other in order to achieve specific and sometimes contradictory goals, are extremely complex. In order to model them well it is necessary to take into account their structure (e.g., who spoke to whom), content (e.g., what was said), and temporal dynamics (e.g., when they spoke).

Traditionally, social processes have been studied by analyzing declared structural links in observed networks, such as "friendships" [1], paper citations [2], and bill co-sponsorships [3]. However, for many domains, explicitly stated links do not exist, are unreliable, or fail to reflect the behaviors of interest. In these situations, the observed *dynamics* of interactions can instead be used as a proxy by which structural properties can be inferred, in order to understand and predict social behavior.

In this paper, we focus on inferring *influence*—a latent attribute that lies at the heart of many social interactions. Since influence is seldom made explicit, it must be inferred from other information. We argue that two forms of reciprocal social behavior can be used to infer and understand influence: conversational turn-taking [4] and linguistic accommodation [5]. Conversational turn-taking,

*Current affiliation: Google DeepMind
†Additional affiliation: UMass Amherst

or "who speaks next," is a fundamental building block of social interactions and can reveal subtle information about power and influence relationships. Similarly, there is a substantial body of work within sociolinguistics about linguistic accommodation, indicating that when two individuals interact, either orally or in writing, the use of a word by one individual will increase the other individual's probability of subsequently using that word. Furthermore, the extent of this increase depends on power differences: the language used by less powerful individuals will drift further so as to more closely resemble or "accommodate" the language used by more powerful individuals.

We present the Bayesian Echo Chamber—a new Bayesian model that jointly models the temporal dynamics and linguistic content of social interactions. This model combines multivariate Hawkes processes [6]—a class of self- and mutually exciting doubly stochastic point processes—with ideas from language modeling. We build on the work of Blundell et al. [7] to model conversational turn-taking, as well as introducing a new linguistic component that models the mutual excitation of words in social interactions. This component, which is based on a discrete analog of Hawkes processes, enables our model to capture linguistic accommodation. We use latent influence variables to provide unique, fine-grained information about who influences whom both temporally and linguistically.

In the next section, we give a brief summary of related work. Then, in section 3, we present the mathematical details of the Bayesian Echo Chamber. We first outline the temporal component and then describe the linguistic component. In section 4, we give an algorithm for inferring the model's latent variables from real-world data. To validate this algorithm, we provide parameter recovery results on synthetic data. In section 5, we evaluate the effects of separating temporal and linguistic influence, compared to a model variant in which influence is tied across both modalities, so as to determine whether a single notion of influence can yield better predictive performance. Finally, we showcase our model's power by analyzing three sources of real-world data: transcripts of real-world conversations; a set of arguments heard by the US Supreme Court [8], consisting of exchanges between justices and attorneys; and a transcript of the movie "12 Angry Men," in which 12 jurors argue about the outcome of a trial. We provide an analysis of the inferred influence patterns, demonstrating state-of-the-art performance at modeling nuanced social dynamics in group discussion settings.

2 Related Work

Several recently developed probabilistic models use point processes as a foundation for inferring social structures from temporal dynamics [9, 7, 10, 11, 12, 13, 14]. Hawkes processes [6] play a central role in some of these models. For example, Blundell et al. [7] used Hawkes processes to infer latent social groups from reciprocal behavior in email data, while Linderman and Adams [13] used Hawkes processes to study patterns of gang-related homicide in Chicago. Although temporal dynamics can reveal some social structures, others may be more readily evidenced by examining both temporal information and interaction content. In this vein, Danescu-Niculescu-Mizil et al. [15] analyzed discussions among Wikipedians and arguments before the US Supreme Court to uncover power differences, while Gerrish and Blei [16] took a language-based approach to measuring scholarly impact and identified influential documents by analyzing changes to thematic content over time. Neither piece of work incorporated a generative model for temporal dynamics. To our knowledge, these two approaches to inferring social structure (modeling temporal dynamics via point processes and modeling interaction content) have not previously been combined in a single Bayesian model.

3 The Bayesian Echo Chamber

In this section, we present our new Bayesian latent variable model—the Bayesian Echo Chamber. The model consists of two components: the first, described in the next section, is a temporal component based on multivariate Hawkes processes that models conversational turn-taking, while the second, described in section 3.2, is a language model that captures linguistic accommodation via a discrete analog of Hawkes processes. These components can be tied together by relating their influence parameters so as to capture a notion of influence that is based upon both temporal and linguistic information. We focus on a broadcast or group discussion setting, in which every utterance is heard by every discussion participant. Our model specifies a probabilistic generative process for a set of utterances $\{\mathcal{W}^{(p)}\}_{p=1}^{P}$ made by P individuals and their corresponding time stamps $\{\mathcal{T}^{(p)}\}_{p=1}^{P}$.

3.1 Temporal Component

Hawkes processes [6], a class of self- and mutually- exciting doubly stochastic point processes, form the foundation of our model's temporal component. A Hawkes process is an inhomogeneous Poisson processes with a stochastic rate function $\lambda(t)$ that depends on events prior to time t. We use $N(\cdot)$ to denote the counting measure of a Hawkes process on $(0,\infty)$, which takes as its argument an interval $[a,b)$ and returns the number of events that occur in $[a,b)$. Its stochastic rate function is

$$\lambda(t) = \lambda_0 + \int_0^{t^-} g(t,u)\,\mathrm{d}N(u) = \lambda_0 + \sum_{n:t_n<t} g(t,t_n), \tag{1}$$

where λ_0 is the base rate of events, t_n is the time stamp of the n^{th} event, and $g(t,u)$ is a non-negative kernel function that specifies the extent to which an event at time $u<t$, increases the instantaneous rate at time t. Typically, $g(t,u)$ is a stationary kernel that decays the influence of events further in the past. Note that $\lambda(t)$ is a *conditional* rate function that depends upon all events prior to time to t.

In our model, the interactions among individuals are captured by a multivariate Hawkes process. Each individual p is associated with a Hawkes process with stochastic rate function

$$\lambda^{(p)}(t) = \lambda_0^{(p)} + \sum_{q\neq p}\int_0^{t^-} g^{(qp)}(t,u)\,\mathrm{d}N^{(q)}(u) = \lambda_0^{(p)} + \sum_{q\neq p}\sum_{n:t_n^{(q)}<t} g^{(qp)}(t,t_n^{(q)}), \tag{2}$$

which is coupled with the Hawkes processes of the other $P-1$ individuals via the respective counting measure $N^{(q)}(\cdot)$. The excitation kernel is chosen to be exponential, i.e., $g^{(qp)}(t,u) = \nu^{(qp)}\exp\left(-(t-u)\,/\,\tau_T^{(p)}\right)$. Parameter $\lambda_0^{(p)}$ is individual p's base rate of utterances, while $\nu^{(qp)}$ controls the degree of instantaneous temporal excitation from individual q to p and $\tau_T^{(p)}$ is a time decay parameter specific to individual p that characterizes how fast excitation decays. In contrast to Blundell et al.'s framework [7], which focused on pairwise interactions and therefore relied upon a bivariate Hawkes process, the multivariate nature of our Hawkes process means that an utterance made by individual p will contribute to the rate function of each of the other $P-1$ individuals[1].

The probability of observing time stamps $\mathcal{T}^{(p)} = \{t_n^{(p)}\}_{n=1}^{N^{(p)}(T)}$ for each p is

$$P(\{\mathcal{T}^{(p)}\}_{p=1}^P \,|\, \{\lambda_0^{(p)}, \{\nu^{(qp)}\}_{q=1}^P, \tau_T^{(p)}\}_{p=1}^P) = \prod_{p=1}^P \left(\exp\left(-\Lambda^{(p)}(T)\right)\prod_{n=1}^{N^{(p)}(T)}\lambda^{(p)}(t_n^{(p)})\right), \tag{3}$$

where $\Lambda^{(p)}(T) = \int_0^T \lambda^{(p)}(t)\,\mathrm{d}t$ is the expected number of utterances in $[0,T)$ [17].

We define the temporal influence of individual q on individual p to be $I_T^{(qp)} = \nu^{(qp)}$. From equation 2, it is evident that a larger value of $\nu^{(qp)}$ results in a higher *instantaneous* rate of utterances for individual p. Therefore, a larger $\nu^{(qp)}$ means a greater temporal influence of individual q on p.

3.2 Linguistic Component

Our model's linguistic component is based on a discrete analog of Hawkes processes and specifies a generative process for utterances $\{\mathcal{W}^{(p)}\}_{p=1}^P$ given their time stamps $\{\mathcal{T}^{(p)}\}_{p=1}^P$. The set of utterances made by individual p is represented by $\mathcal{W}^{(p)} = \{\boldsymbol{w}_n^{(p)}\}_{n=1}^{N^{(p)}(T)}$, where each utterance $\boldsymbol{w}_n^{(p)}$ consists of a set of $L_n^{(p)}$ tokens $\{w_{l,n}^{(p)}\}_{l=1}^{L_n^{(p)}}$. The generative process for each token $w_{l,n}^{(p)}$ is

$$w_{l,n}^{(p)} \sim \text{Cat}\,(\boldsymbol{\phi}_n^{(p)}) \quad \text{and} \quad \boldsymbol{\phi}_n^{(p)} \sim \text{Dir}\,(\alpha^{(p)}, \boldsymbol{B}_n^{(p)}), \tag{4}$$

where V-dimensional probability vector $\boldsymbol{\phi}_n^{(p)}$ is is specific to individual p's n^{th} utterance, $\alpha^{(p)}$ is a scalar concentration parameter and $\boldsymbol{B}_n^{(p)}$ is a V-dimensional normalized base measure satisfying

$$\mathrm{B}_{v,n}^{(p)} \propto \beta_v^{(p)} + \sum_{q\neq p}\rho^{(qp)}\psi_{v,n}^{(qp)} \quad \text{and} \quad \sum_{v=1}^V B_{v,n}^{(p)} = 1 \tag{5}$$

[1] Self-excitation is not permitted; we are interested in the interactions between individuals.

Vector $\boldsymbol{\beta}^{(p)} \in \mathbb{R}_+^V$ characterizes individual p's inherent language usage. Parameter $\rho^{(qp)}$ controls the degree of linguistic excitation or "influence" from individual q to p, while $\boldsymbol{\psi}_n^{(qp)} \in \mathbb{R}_+^V$ is a vector of decayed excitation pseudocounts from q to p just before p makes his or her n^{th} utterance:

$$\psi_{v,n}^{(qp)} = \sum_{m:t_m^{(q)} < t_n^{(p)}} \left(\sum_{l=1}^{L_m^{(q)}} \mathbf{1}(w_{l,m}^{(q)} = v) \right) \exp\left(-\frac{t_n^{(p)} - t_m^{(q)}}{\tau_L^{(p)}} \right). \tag{6}$$

The inner sum corresponds to the number tokens of type v in individual q's m^{th} utterance. Parameter $\tau_L^{(p)}$ is a linguistic time decay specific to individual p. Thanks to Dirichlet–multinomial conjugacy, the probability vectors $\{\{\boldsymbol{\phi}_n^{(p)}\}_{n=1}^{N^{(p)}(T)}\}_{p=1}^P$ can be analytically integrated out so as to yield

$$P(\{\mathcal{W}^{(p)}\}_{p=1}^P \,|\, \{\mathcal{T}^{(p)}, \alpha^{(p)}, \boldsymbol{\beta}^{(p)}, \tau_L^{(p)}, \{\rho^{(qp)}\}_{q=1}^P\}_{p=1}^P) = \tag{7}$$

$$\prod_{p=1}^P \prod_{n=1}^{N^{(p)}(T)} \prod_{l=1}^{L_n^{(p)}} \frac{\sum_{l'=1}^{l-1} \mathbf{1}(w_{l',n}^{(p)} = w_{l,n}^{(p)}) + \alpha^{(p)} \mathrm{B}_{w_{l,n}^{(p)},n}^{(p)}}{l - 1 + \alpha^{(p)}}. \tag{8}$$

We define the linguistic influence of individual q on individual p to be $I_L^{(qp)} = \rho^{(qp)}$. A larger value of $\rho^{(qp)}$ reflects a greater influence of q's utterances' content on p's subsequent word usage.

3.3 Shared Influence

The influence and linguistic influence can be tied together by setting $\rho^{(qp)} = r\nu^{(qp)}$, where r is a scaling factor. Intuitively, $\nu^{(qp)}$ has units of utterances per unit time, while $\rho^{(qp)}$ has units of tokens per utterance. Since each utterance involves some number of tokens, these parameters are linearly related. Tying the influence parameters in this way provides the model with the capacity to capture a global notion of influence that is based upon both temporal dynamics and linguistic information. In section 5, we provide an empirical comparison of this tied model variant with the untied version.

4 Inference

For real-world data, the utterances $\{\mathcal{W}^{(p)}\}_{p=1}^P$ and their time stamps $\{\mathcal{T}^{(p)}\}_{p=1}^P$ are observed, while parameters $\{\lambda_0^{(p)}, \{\nu^{(qp)}\}_{q\neq p}, \tau_T^{(p)}, \tau_L^{(p)}, \alpha^{(p)}, \boldsymbol{\beta}^{(p)},\}_{p=1}^P$ and, depending on the model variant, r or $\{\{\rho^{(qp)}\}_{q\neq p}\}_{p=1}^P$ are unobserved. We therefore sample typical values for these latent variables from their posterior distribution given $\{\mathcal{W}^{(p)}, \mathcal{T}^{(p)}\}_{p=1}^P$ using a slice-within-Gibbs algorithm [18] that operates by sequentially resampling the value of each variable from its conditional posterior. Since $\{\lambda_0^{(p)}\}_{p=1}^P$, $\{\{\nu^{(qp)}\}_{q\neq p}\}_{p=1}^P$, $\{\tau_T^{(p)}\}_{p=1}^P$, and, in the case of the untied variant, $\{\{\rho^{(qp)}\}_{q\neq p}\}_{p=1}^P$ and $\{\tau_L^{(p)}\}_{p=1}^P$, are highly correlated, slice-within-Gibbs can get stuck in local modes. We therefore initialize these variables using values obtained via sequential least squares programming [19].

To ensure stationarity, we employ the stationarity condition of Bremaud and Massouli [20]. If $\boldsymbol{M}$ is a $P \times P$ matrix where $M_{qp} = \int_u^\infty \left| g^{(qp)}(t,u) \right| \, \mathrm{d}t = \nu^{(qp)} \tau_T^{(p)}$, this condition requires that the spectral radius of $\boldsymbol{M}$ be strictly less than one. Since the spectral radius of $\boldsymbol{M}$ is upper-bounded by any matrix norm, the condition may be enforced by requiring that $\|\boldsymbol{M}\| < 1$ for some norm $\|\cdot\|$. For simplicity, we choose the maximum absolute column sum norm $\|\cdot\|_1$, which leads to $\|\boldsymbol{M}\|_1 = \max_{p=1,\ldots,P} \sum_{q\neq p} |\nu^{(qp)} \tau_T^{(p)}| = \max_{p=1,\ldots,P} \tau_T^{(p)} \sum_{q\neq p} \nu^{(qp)} < 1$. Rearrangingyields the following upper bounds for $\{\{\nu^{(qp)}\}_{q\neq p}\}_{p=1}^P$ and and $\{\tau_T^{(p)}\}_{p=1}^P$: $\nu^{(qp)} < \frac{1}{\tau_T^{(p)}} - \sum_{u\neq q, u\neq p} \nu^{(up)}$ and $\tau_T^{(p)} < \frac{1}{\sum_{q\neq p} \nu^{(qp)}}$. We enforce these upper bounds, as well as the lower bounds of zero for these variables and $\{\lambda_0^{(p)}, \{\rho^{(qp)}\}_{q\neq p}, \tau_L^{(p)}\}_{p=1}^P$ throughout inference. Finally, we assume that $\{\alpha^{(p)}\}_{p=1}^P \sim \mathrm{Gamma}\,(10, 10)$, $\{\{\beta_v^{(p)}\}_{v=1}^V\}_{p=1}^P \sim \mathrm{Gamma}\,(10, 20)$, and $r \sim \mathrm{Gamma}\,(5, 4)$.

Data set	No. individuals	No. utterances	No. tokens
City meeting	2	65	2,270
Family discussion	8	396	2,145
University lecture	5	148	3,099
L&G v. Texas	6	201	7,734
DC v. Heller	10	364	13,746
Citzens United v. FEC	10	344	11,446
"12 Angry Men"	12	312	5,713

Table 1: Salient characteristics of each real-world data set.

5 Experiments

In this section we first validate our inference algorithm using synthetic data, and then compare the log probabilities obtained using our tied and untied model variants on several real-world sources of data: conversation transcripts from the Santa Barbara Corpus of Spoken American English, transcripts of oral arguments heard by the US Supreme Court, and the transcript of the 1957 movie "12 Angry Men." Finally, we demonstrate our model's ability to discover influential participants and latent social structures in Supreme Court arguments and the transcript of "12 Angry Men."

We selected the three conversations with the largest numbers of participants from the Santa Barbara Corpus of Spoken American English [8]. These conversations consist of a city meeting, a family discussion, and a university lecture. As a result, they capture the natural back-and-forth of many real-world conversations, both in terms of their temporal dynamics and their linguistic content.

The US Supreme Court consists of a chief justice and eight associate justices. Each oral argument heard by the Court therefore involves up to nine justices (some may recuse themselves) plus attorneys representing the petitioner and the respondent. The format of each argument is formulaic: the attorneys for each party have 30 minutes to present their argument, with those representing the petitioner speaking first. Justices routinely interrupt these presentations to make comments or ask questions of the attorneys. Sometimes additional attorneys, known as *amicae curae*, also present arguments in support of either the petitioner or the respondent. We used the time stamped transcripts from three controversial Supreme Court cases [8]: Lawrence and Garner v. Texas, District of Columbia v. Heller, and Citizens United v. Federal Election Commission (re-argument only).

"12 Angry Men" is movie about a jury's deliberations regarding the guilt or acquittal of a defendant. Unlike Supreme Court arguments, the dialog is informal and seemingly natural, rather than formal and prescribed. The movie is unique in its limited cast of only 12 individuals and the fact that it is set almost entirely in one room. These qualities, combined with the fact that the movie explicitly focuses on discussion-based consensus-building in a group setting, make its time stamped transcript an ideal data set for exploring the strengths of our model. We generated an appropriate transcript from the movie subtitles by hand-labeling the speaker of each utterance in the first half of the movie.

For all data sets, we concatenated concurrent utterances by the same individual, discarded contributions from individuals with fewer than ten (post-concatenation) utterances, and rescaled all time stamps to lie in the interval $(0, 1]$. We did not remove stopwords, since they can carry important information about power and influence [15], but we did discard all but the 300 most frequent stemmed word types. The salient characteristics of each data set, after preprocessing, are shown in table 1.

5.1 Parameter Recovery on Synthetic Data

We used an exact simulation algorithm for Hawkes processes [21] to generate synthetic data (consisting of 3 individuals, 628 utterances, and 62,800 tokens spanning 20 word types) from each of our model variants and then used the inference algorithm described in the previous section to infer typical values for parameters $\{\lambda_0^{(p)}, \{\nu^{(qp)}\}_{q\neq}^P, \tau_T^{(p)}, \alpha^{(p)}, \boldsymbol{\beta}^{(p)}, \tau_L^{(p)}\}_{p=1}^P$ and, depending on the model variant, either δ or $\{\{\rho^{(qp)}\}_{q\neq p}\}_{p=1}^P$. We compared these inferred parameters values to the values used to generate the synthetic data. Figure 1 shows the true and inferred parameter values for the tied model variant. The algorithm does a good job of recovering the true parameter values. The true parameter values for the untied variant are recovered similarly well; we omit them to save space.

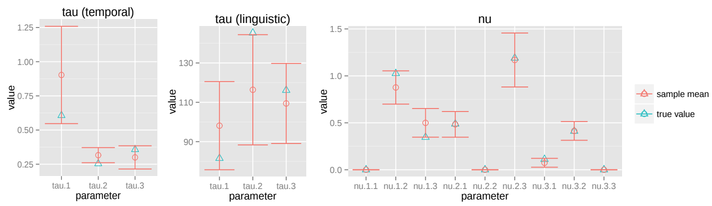

Figure 1: Parameter recovery on synthetic data for the tied model variant.

Data set	Tied	Untied	No influence	DTM+Poisson
City meeting	**-764.88**$\pm$0.07	-769.42$\pm$0.07	-797.75$\pm$0.06	-838.13
Family discussion	-1073.55$\pm$0.21	**-1062.42**$\pm$0.21	-1121.67$\pm$0.11	-1046.72
University lecture	-464.47$\pm$0.15	**-461.07**$\pm$0.16	-523.05$\pm$0.07	-469.72
L&G v. Texas	-5507.11$\pm$0.15	**-5502.87**$\pm$0.15	-5721.19$\pm$0.18	-5542.39
DC v. Heller	-6321.30$\pm$0.16	**-6303.55**$\pm$0.15	-6572.17$\pm$0.26	-6461.25
Citizens United v. FEC	-4795.24$\pm$0.18	**-4777.96**$\pm$0.17	-5239.52$\pm$0.23	-4812.22
"12 Angry Men"	-4014.56$\pm$0.24	**-3987.20**$\pm$0.23	-4615.50$\pm$0.19	-4107.38

Table 2: Log probabilities on held-out data obtained using both model variants and two baselines–one without influence, and one based upon the dynamic topic model and Poisson processes.

5.2 Log Probabilities of Held-Out Data

In order to quantify the effect of tying $\rho^{(qp)}$ and $\nu^{(qp)}$, we split each data set into a "training" split $\mathcal{D}_{\text{train}} = \{\mathcal{W}^{(p)}_{\text{train}}, \mathcal{T}^{(p)}_{\text{train}}\}_{p=1}^{P}$ and a "held-out" or "test" split $\mathcal{D}_{\text{test}} = \{\mathcal{W}^{(p)}_{\text{test}}, \mathcal{T}^{(p)}_{\text{test}}\}_{p=1}^{P}$. We formed each training split by selecting only those utterances that occurred before time t^*, where t^* was chosen so as to yield a 90%–10% training–testing split. For each model variant (i.e., tied or untied) and data set, we then inferred typical values for that model variant's parameters by drawing samples from the posterior distribution over parameters, conditioned on $\mathcal{D}_{\text{train}}$, using our inference algorithm.

Letting $\mathcal{H}^{(s)}$ denote a single set of sampled parameters augmented with the fixed hyperparameters described in section 4, we can form an approximate lower bound on the log probability of held-out utterances $\mathcal{D}_{\text{test}}$ given training utterances $\mathcal{D}_{\text{test}}$ via the concave nature of the log function, $\log\left(P(\mathcal{D}_{\text{test}} \,|\, \mathcal{D}_{\text{train}})\right) \geq \frac{1}{S}\sum_{s=1} \log\left(P(\mathcal{D}_{\text{test}} \,|\, \mathcal{H}^{(s)})\right)$. Approximate held-out log probabilities for each model variant (tied or untied) and each data set (obtained using $S = 2000$ samples after 1000 burn-in iterations) are given in table 2. We also provide held-out log probabilities for two baseline models: one consisting of our model with all influence parameters set to zero, and one consisting of the dynamic topic model [22] plus a Poisson process. For almost all data sets, our untied model variant, which models temporal influence and linguistic influence separately, outperformed our tied model variant. These results imply that for many real-world settings there are fundamental differences between the latent attributes captured by the temporal and linguistic influence parameters.

5.3 US Supreme Court

Temporal influence parameters $\{\{I_T^{(qp)}\}_{q\neq p}^{P}\}_{p=1}^{P}$ inferred from the District of Columbia v. Heller transcript are shown in figure 2a, while the total temporal influences to and from each individual (computed by summing the columns and rows of the influence matrix) are shown in figure 2c. The justices present for this case were Alito, Breyer, Ginsburg, Kennedy, Roberts, Scalia, Stevens, Souter, and Thomas, while the attorneys were Dellinger (representing the petitioner), Gura (representing the respondent), and Clement (as amicae curae, supporting the petitioner). Ultimately, Alito, Kennedy, Roberts, Scalia, and Thomas (the majority) sided with the respondent, while Breyer, Ginsburg, and Stevens (the minority) sided with the petitioner. Neither Alito or Thomas spoke ten or more utterances, so they are not present in figure 2. The most striking pattern in figure 2a is that the three attorneys received much more temporal influence from the justices than vice versa. This reflects not only the status difference between these two groups, but also the structure of the argument: attorneys present arguments, while justices interrupt their presentations to make comments or ask questions. Inferred linguistic influence parameters $\{\{I_L^{(qp)}\}_{q\neq p}^{P}\}_{p=1}^{P}$ and total linguistic influences

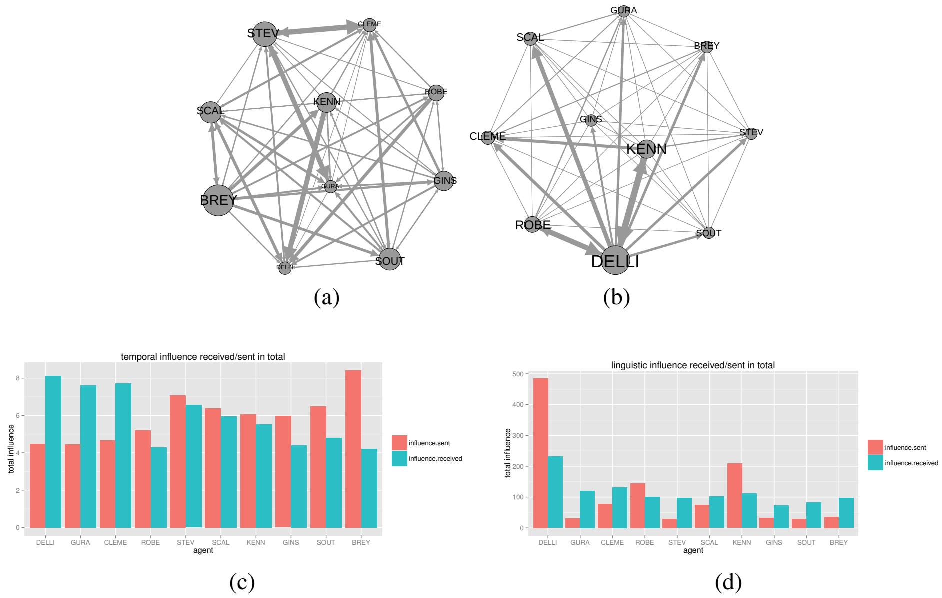

Figure 2: (a) Graph of inferred temporal influence parameters for District of Columbia v. Heller. Each node corresponds to a justice or attorney and the width of each edge corresponds to the associated temporal influence $I_T^{(qp)}$. (b) Graph of inferred linguistic influence parameters. (c) Total inferred temporal influence to and from each individual. (d) Total inferred linguistic influence.

to and from each individual are shown in figures 2b and 2d, respectively. The influence parameters are very sparse, with Dellinger exerting the most influence. Dellinger represented the petitioner and therefore presented his argument first, thereby framing the discussion. The justices with the most linguistic influence were Kennedy and Roberts, both of whom ultimately supported the respondent.

The other two cases, Lawrence and Garner v. Texas and Citizens United v. Federal Election Commission exhibited remarkably similar temporal and linguistic influence patterns to District of Columbia v. Heller: the attorneys, particularly those explicitly representing the petitioner or respondent, received more temporal influence from justices than vice versa, while the linguistic influence matrix is dominated by the petitioner's attorney and a small number of justices who ultimately sided with the respondent, typically the more ideologically moderate ones. The temporal influence patterns reflect the hierarchical structure of the Supreme Court, in which the justices routinely interrupt the attorneys' presentations, inciting a fast response. In contrast, the linguistic influence patterns primarily reflect the nature of Supreme Court arguments: the petitioner's attorney always presents first and therefore has the opportunity to frame the discussion. Some of the justices who ultimately sided with the respondent also hold significant linguistic influence. We speculate that this is because those are the justices whose opinions the petitioner's attorney is trying the hardest to sway.

5.4 12 Angry Men

A graph of the inferred temporal influence parameters $\{\{I_T^{(qp)}\}_{q \neq p}^P\}_{p=1}^P$ is shown in figure 3a, while the total temporal influence to and from each individual is shown in figure 3c. The most significant pattern in these figures is that four jurors have much more influence over others than others do over them (Juror 2, Juror 5, Juror 9, and Juror 11) while the reverse is true for Jurors 3 and 8. Juror 8 is the protagonist of the movie, and initially casts the only "not guilty" vote. Ultimately, the other jurors to change their votes to match his. As he argues with the other jurors, he must respond to their questions and defend his position. Juror 3, is the last one to change his vote. The four jurors who exert more influence than they receive are the first four jurors to change their votes. Jurors 5 and 11, who exert the most influence, are verbose, while Jurors 2 and 9 are comparatively taciturn.

The inferred linguistic influence parameters $\{\{I_L^{(qp)}\}_{q \neq p}^P\}_{p=1}^P$ are shown in figure 3b, while the total linguistic influences to and from each individual are shown in figure 3d. Here, the most significant

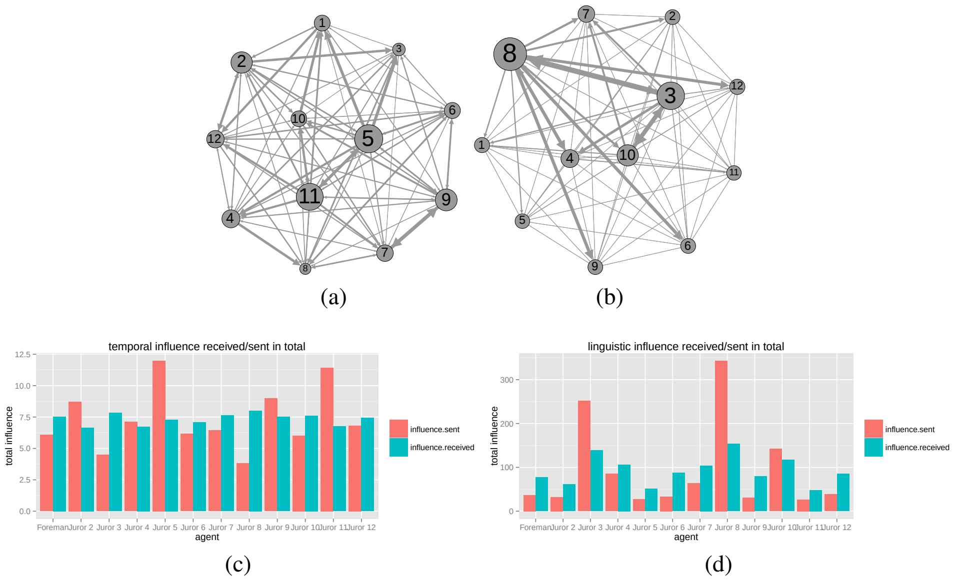

Figure 3: (a) Graph of inferred temporal influence parameters for the "12 Angry Men" data set. Each node corresponds to a character in the film, the width of each edge corresponds to the associated temporal influence $I_T^{(qp)}$. (b) Inferred linguistic influence parameters. (c) Total inferred temporal influence to and from each individual. (d) Total inferred linguistic influence.

pattern is that three individuals have much more linguistic influence over others than others do over them: Juror 8, Juror 3, and to a lesser extent Juror 10. It is unsurprising that Juror 8, the first to vote "not guilty" should dominate the discussion content. Similarly, Juror 3, the last to change his vote to match Juror 8's, is most invested in discussing defendant's supposed guilt. Furthermore, according to figure 3d, Juror 8 has the least linguistic influence over Juror 3. Lastly, Juror 10, along with Juror 4, is one of the last three jurors to change his vote. However, unlike Juror 4 (who stands out only marginally in figure 3d and, according to figure 3d, has less influence over others than others do over him), Juror 10 is argumentative in the lead up to his change of mind. Overall, the consistency of the inferred influence parameters (both temporal and linguistic) with the narrative of the movie, suggests that our model is indeed capable of uncovering latent influence relationships from interaction content. This is also reflected in our model's ability to predict held-out data well.

6 Discussion

We introduced a new Bayesian latent variable model, the Bayesian Echo Chamber, which discovers temporal and linguistic influence patterns, via conversational turn-taking and linguistic accommodation. We investigated whether these two types of influence are better modeled with a single set of influence parameters and discovered that in most cases temporal and linguistic influence capture different aspects of social behavior and are best represented separately. We demonstrated that the Bayesian Echo Chamber can jointly predict both the time stamps and content of future interactions, and outperforms baseline models in doing so. Finally, we showcased the model's ability to discover influential individuals in group discussions using data from the US Supreme Court and the movie "12 Angry Men," and provided extensive discussion of the resultant influence patterns. Avenues for future work include evaluating linguistic accommodation separately on function and content words, exploring different ways of tying model parameters so as to model different social structures, and modeling data sets for which other Hawkes process kernel functions may be more appropriate.

References

[1] L. Backstrom, D. Huttenlocher, J. Kleinberg, and X. Lan. Group formation in large social networks: Membership, growth, and evolution. In *Proceedings of the 12th ACM SIGKDD*

International Conference on Knowledge Discovery and Data Mining, 2006.

[2] D. J. de Solla Price. Networks of scientific papers. *Science*, 149(3683):510–515, 1965.

[3] J. H. Fowler. Legislative cosponsorship networks in the US House and Senate. *Social Networks*, 28:454–465, 2006.

[4] H. Sacks. *Lectures on Conversation*, volume 1. Blackwell Publishing, 1995.

[5] R. West and L. Turner. *Introducting Communication Theory: Analysis and Applications*. McGraw Hill, 2010.

[6] A. G. Hawkes. Point spectra of some self-exciting and mutually-exciting point processes. *Journal of the Royal Statistical Society: Series B (Methodology)*, 58:83–90, 1971.

[7] C. Blundell, K. A. Heller, and J. Beck. Modelling reciprocating relationships with Hawkes processes. In *Advances In Neural Information Processing Systems*, 2012.

[8] B. MacWhinney. The TalkBank project. *Creating and Digitizing Language Corpora: Synchronic Databases*, 2007.

[9] A. Simma and M. I. Jordan. Modeling events with cascades of Poisson processes. *Proceedings of Uncertainty in Artificial Intelligence*, 2010.

[10] P. O. Perry and P. J. Wolfe. Point process modelling for directed interaction networks. *Journal of the Royal Statistical Society: Series B (Methodology)*, 75(5):821–849, 2013.

[11] T. Iwata, A. Shah, and Z. Ghahramani. Discovering latent influence in online social activities via shared cascade poisson processes. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 266–274, 2013.

[12] C. DuBois, C. Butts, and P. Smyth. Stochastic blockmodeling of relational event dynamics. In *Proceedings of the Sixteenth International Conference on Artificial Intelligence and Statistics*, pages 238–246, 2013.

[13] S. W. Linderman and R. P. Adams. Discovering latent network structure in point process data. *arXiv preprint 1402.0914*, 2014.

[14] K. Zhou, H. Zha, and L. Song. Learning triggering kernels for multi-dimensional Hawkes processes. In *Proceedings of the 30th International Conference on Machine Learning*, pages 1301–1309, 2013.

[15] C. Danescu-Niculescu-Mizil, L. Lee, B. Pang, and J. Kleinberg. Echoes of power: Language effects and power differences in social interaction. In *Proceedings of the 21st International Conference on World Wide Web*, pages 699–708, 2012.

[16] S. M. Gerrish and D. M. Blei. A language-based approach to measuring scholarly impact. In *Proceedings of the 26th International Conference on Machine Learning*, 2010.

[17] D. J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes*. Springer, 1988.

[18] R. M. Neal. Slice sampling. *Annals of Statistics*, 31(3):705–767, 2003.

[19] D. Kraft. A software package for sequential quadratic programming. Technical Report DFVLR-FB 88-22, DLR German Aerospace Center, Institute for Flight Mechanics, 1988.

[20] P. Bremaud and L. Massouli. Stability of nonlinear Hawkes processes. *The Annals of Probability*, pages 1563–1588, 1996.

[21] A. Dassios and H. Zhao. Exact simulation of Hawkes process with exponentially decaying intensity. *Electronic Communications in Probability*, 18(62), 2013.

[22] D. M. Blei and J. D. Lafferty. Dynamic topic models. In *Proceedings of the 23rd International Conference on Machine Learning*, 2006.