Proceedings of the Nineteenth International AAAI Conference on Web and Social Media (ICWSM 2025)

Variational Inference of Parameters in Opinion Dynamics Models

Jacopo Lenti[1,2], **Fabrizio Silvestri**[1], **Gianmarco De Francisci Morales**[2]
[1]Sapienza University, P.le A. Moro 5, Rome, Italy, 00185
[2]CENTAI, Corso Inghilterra, 3, Turin, Italy, 10138
jcp.lenti@gmail.com, fsilvestri@diag.uniroma1.it, gdfm@acm.org

Abstract

Modeling human behavior through the lens of online social networks presents both a significant opportunity and a challenge for understanding complex social phenomena, such as misinformation spread, opinion formation and polarization. While agent-based models (ABMs) are widely used for studying these social phenomena, parameter estimation remains a challenge, often relying on costly simulation-based heuristics. This work uses variational inference to estimate the parameters of an opinion dynamics ABM by transforming the estimation problem into an optimization task that can be solved directly.

Our proposal relies on probabilistic generative ABMs (PGABMs): we start by synthesizing a probabilistic generative model from the ABM rules. Then, we transform the inference process into an optimization problem suitable for automatic differentiation. In particular, we use the Gumbel-Softmax reparameterization for categorical agent attributes and Stochastic Variational Inference for parameter estimation. Moreover, we explore the trade-offs of using variational distributions with different complexities: Normal distributions and Normalizing Flows. We validate our method on a bounded confidence model with agent roles, by estimating both macroscopic (bounded confidence intervals and backfire thresholds) and microscopic (200 categorical agent-level roles) parameters more accurately than simulation-based and MCMC methods.

Introduction

The wide availability of data from social media platforms created unprecedented opportunities to study human behavior and its social implications (Peralta, Kertész, and Iñiguez 2022). For this reason, the modeling of human behaviors in the Web has become essential to understand the modern societies (Chen, Xiao, and Kumar 2023).

A vaste portion of Computational Social Science literature employed Agent-Based Models (ABMs) to investigate the online dynamics driving misinformation spread (Cinelli et al. 2021), formation of echo chambers (Cinelli et al. 2020) and health communication (Prieto Curiel and González Ramírez 2021; Sobkowicz and Sobkowicz 2021). ABMs are computational frameworks aimed to describe human behaviors in social systems and the consequent emergence of complex phenomena, by simulating the actions and interactions of autonomous agents within a social network (Grazzini, Richiardi, and Tsionas 2017). In ABMs, the agents possess individual attributes and act following predefined behavioral rules, that are simulated to generate outcomes that are validated against real-world data. While ABMs offer great flexibility and a clear representation of the causal mechanisms that govern the system, they pose significant challenges in the inference phase. The obstacles come from several factors, such as the considerable computational effort required to generate data (a complete simulation of the model) and the high dimensionality of microscopic parameters, which scale with the number of agents. Additionally, the heterogeneity of the agents and the diversity of the rules in the ABM literature hinder the tractability of these models (Platt 2020).

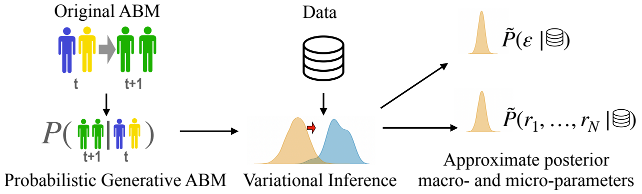

Figure 1: First, we translate the agent-based model into a probabilistic generative agent-based model. Then, we apply variational inference to get an approximate posterior of the target parameters within a given dataset.

The current practice for parameter estimation in ABMs mostly focuses on simulation-based methods. These methods compute parameter estimates by running a large amount of computationally expensive simulations of the model with different values of the parameters and comparing the observed and simulated traces of the system state through some summary statistics (a process called *calibration*) (Fagiolo et al. 2019; Lux and Zwinkels 2018). These summary statistics cause an inevitable loss of information at a granular scale. When the parameters are high-dimensional, an exhaustive exploration of the parameter space is not feasible. For this reason, simulation-based approaches have been mostly used to estimate low-dimensional parameters while ignoring the microscopic parameters whose dimensionality scales with the number of agents (Srikrishnan and Keller 2021; Fagiolo et al. 2019). Approximate Bayesian Computation (ABC) is one popular simulation-based approach (Carrella 2021).

A recent line of work has suggested tackling the parameter estimation problem in ABMs radically differently: via a likelihood-based approach (Monti, De Francisci Morales, and Bonchi 2020; Monti et al. 2022; Lenti, Monti, and De Francisci Morales 2024). These works propose a paradigm shift by translating the ABMs into probabilistic generative models, called Probabilistic Generative ABMs (PGABMs). PGABMs explicitly define the latent and observed variables of the system according to the available data and describe the conditional probabilities connecting them from the rules of the ABMs. In this way, it is possible to derive the likelihood function of the latent variables. As a result, this approach is more principled, more accurate, and faster than simulation-based alternatives (Lenti, Monti, and De Francisci Morales 2024). Nevertheless, the likelihood-based approach requires a non-trivial analytical derivation of the likelihood function. In addition, handling categorical variables is still an open challenge: for instance, a maximum-likelihood approach in our setting would require exploring an exponentially large combination of agents' parameter values (up 2^{200}). In this work we circumvent the need to write an explicit derivation of the likelihood by using Variational Inference (VI). VI assumes a tractable, parametric, functional form for the approximation of the posterior distribution of the parameters to be estimated. Then, it directly minimizes the divergence of the approximated distribution from the real posterior. In so doing, our VI approach can address models with intractable likelihood functions (e.g., with a large number of categorical variables). Figure 1 depicts the proposed pipeline.

To demonstrate the viability of our approach, we apply it to a bounded confidence model where each agent is parameterized by a categorical role (Weisbuch et al. 2002). In our experiments, the VI-based approach is 3× more accurate in estimating agents' roles than an MCMC baseline and almost 8× more accurate than a simulation-based ABC. In addition, VI is an order of magnitude faster than the alternatives.

Opinion Dynamics Model

We focus on a model based on bounded confidence model with backfire effect (Jager and Amblard 2005). Each agent has an opinion X_u in $[0, 1]$. At each time step, an agent interacts with one of its neighbors. In the original model, if the opinions of the two interacting agents are closer than the bounded confidence interval (ε^+) they have a positive interaction ($s = 1$), and their opinions converge by a convergence rate (μ^+). Conversely, if their opinions are further than a backfire threshold (ε^-), then they have a negative interaction ($s = -1$), and their opinions diverge by a divergence rate (μ^-). In the other cases, their opinions remain unchanged ($s = 0$). We extend the model to represent a social context with Leaders and Followers. Each agent has a role (r_u) which can be either *leader* ($r_u = L$) or *follower* ($r_u = F$). The behavior of Leaders and Followers is driven, respectively, by ε^+_L and ε^-_L, and ε^+_F and ε^-_F. We assume that followers are more inclined to change their opinions, and thus $\varepsilon^+_F \geq \varepsilon^+_L$, $\varepsilon^-_F \leq \varepsilon^-_L$, $\mu^+_F \geq \mu^+_L$, and $\mu^-_F \geq \mu^-_L$. For conciseness, we define $\varepsilon = (\varepsilon^+_F, \varepsilon^+_L, \varepsilon^-_F, \varepsilon^-_L)$, $\mu = (\mu^+_F, \mu^+_L, \mu^-_F, \mu^-_L)$ and r the vector of agents' roles.

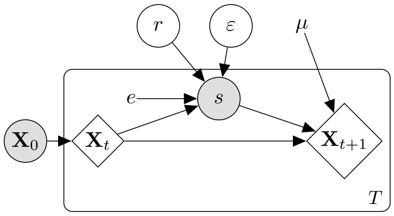

Figure 2: Probabilistic Graphical Model associated with the Bounded Confidence Model with backfire effect and latent roles. Circles represent stochastic variables, diamonds deterministic variables, and letters are given parameters. Shaded variables are observed and white ones are latent.

Inference

We aim to estimate the latent variables and unknown parameters of the opinion dynamics model. Firstly, we translate the ABM into the corresponding PGABM. Secondly, we ensure the differentiability of the model by relaxing the discrete distributions. Finally, we apply Variational Inference (VI) to estimate the target variables.

To translate the ABM into its PGABM counterpart, we define the ABM rules as conditional probabilities. We model separately the probability of having a positive and a negative interaction. Let s^+_j and s^-_j the outcomes of the interaction j,

$$P(s^+_j = 1 \mid x_t, e_j, \varepsilon, r_v) = \sigma(\rho \cdot (\varepsilon^+_{r_v} - |\Delta x^t_{uv}|)) \quad (1)$$

$$P(s^-_j = 1 \mid x_t, e_j, \varepsilon, r_v) = \sigma(-\rho \cdot (\varepsilon^-_{r_v} - |\Delta x^t_{uv}|)), \quad (2)$$

where $\sigma(\cdot)$ is the sigmoid function and ρ is its steepness. In this way, we have a probabilistic model where we observe $s^+ = 1$ with a high probability if $|\Delta x^t_{uv}| < \varepsilon^+_{r_v}$, and $s^- = 1$ if $|\Delta x^t_{uv}| \geq \varepsilon^-_{r_v}$. This process ensures that any sample from the parameter space can be associated with a non-zero probability of having generated the data.

Then, we use the probabilistic graphical model (Figure 2) to capture the repeated patterns and the conditional independencies between the random variables. To this end, we identify the observed variables, the latent variables, and the PGM parameters, together with the relationships that connect them. In our setting, we assume that the pairs of interacting agents (e), the outcomes of the interactions (s^+, s^-), the initial opinions (X_0), and the convergence and divergence rates (μ) are observed. X_t is a deterministic variable, and it is computable from previous observations. Note that μ and e only affect the observed variables, so we consider them as parameters of the PGM. Instead, ε and the agent roles r are latent. In particular, we consider r as a micro-parameter, as it is a vector of size N, while ε has dimension 4. So, overall, we need to estimate $N + 4$ latent variables.

To estimate the variables, we adopt a VI routine. VI is an alternative to MCMC algorithms for approximating a target density via an optimization step (Blei, Kucukelbir, and McAuliffe 2017). VI approximates the target posterior $p(\theta \mid y)$ by using a more tractable family of densities $q_\lambda(\theta)$, called the *variational distribution*, indexed by the variational parameter λ. The optimal variational parameter λ^* is the one that minimizes the KL-divergence of $q_\lambda(\theta)$ from $p(\theta \mid y)$. In practice, the search for this parameter is done by maximizing the evidence lower bound (ELBO), $\mathcal{L}(\lambda) = \mathbb{E}_q[\log p(y, \theta) - q_\lambda(\theta)]$, which is equivalent to minimizing the KL-divergence of $q_\lambda(\theta)$ from $p(\theta \mid y)$.

However, maximizing the ELBO via gradient ascent is tractable only for continuous and differentiable variables. Categorical variables pose a great challenge, both because of their discrete nature and because there is no inherent ordering within their support. The Gumbel-Softmax reparameterization offers a viable solution for these cases (Maddison, Mnih, and Teh 2016; Jang, Gu, and Poole 2016). Rather than sampling a variable from a categorical distribution with probabilities $(\pi_1, \dots, \pi_k)$, we sample a smooth approximation of this distribution. If $k = 2$, the Gumbel-Softmax reparameterizes a Bernoulli distribution. The Gumbel-Softmax relaxation allows us to compute the gradient of the ELBO in a step of automatic differentiation (AD). At this point, we have transformed the inference into an optimization task that we can tackle within an AD framework. Using Adam optimizer (Kingma and Ba 2014), we estimate both the microscopic and macroscopic target parameters.

Experiments

We compare 4 different methods to infer the target latent variables. Two VI methods, representing our proposal, an MCMC method and Approximate Bayesian Computation (ABC), a popular Bayesian simulation-based approach[1].

The choice of the variational distribution q_λ is a critical step in the VI process, as it encodes the functional form of the approximation of $p(\theta \mid y)$. Thanks to its simplicity, the family of normal distributions is a common choice. Another more flexible solution are normalizing flows (NFs) (Rezende and Mohamed 2015). Due to their expressiveness, NFs represent arbitrarily complex distributions, which we can efficiently sample from and compute the ELBO of (Papamakarios et al. 2021). We adopt Stochastic Variational Inference (SVI) to maximize the ELBO (Hoffman et al. 2013).

An alternative approach to estimate the latent variables of the same PGM is sampling-based inference via an MCMC algorithm. We use the No-U Turn Sampler (NUTS), an adaptive Hamiltonian Monte Carlo method that is known for its high efficiency and flexibility (Hoffman, Gelman et al. 2014).

On the other side, ABC is a likelihood-free simulation-based method (Csilléry et al. 2010). ABC samples a set of parameters from a prior distribution, and it runs the entire simulation of the opinion dynamics model for each sample. We use the number of positive interactions as summary statistics for comparing the simulated and observed data trajectories.

We run a grid of 912 experiments, where we vary $N \in \{50, 100, 200, 400\}$, $T \in \{128, 512, 2048, 8192\}$ and leaders proportion $\in \{0.01, 0.02, 0.04, 0.1, 0.2\}$. Each experiment samples ε_F^+ and ε_L^+ in $\{0.05, 0.15, 0.25, 0.35, 0.45\}$, such that $\varepsilon_F^+ \geq \varepsilon_L^+$, and ε_F^- and ε_L^- in $\{0.55, 0.65, 0.75, 0.85, 0.95\}$, such that $\varepsilon_F^+ \leq \varepsilon_L^+$. For each combination of parameters, we simulate the ABM and then estimate the target parameters with the four methods. Since the methods under scrutiny are all Bayesian, the parameter estimates are distributions. To ease understanding of the results, we compare the average values of 200 samples from each posterior distribution.

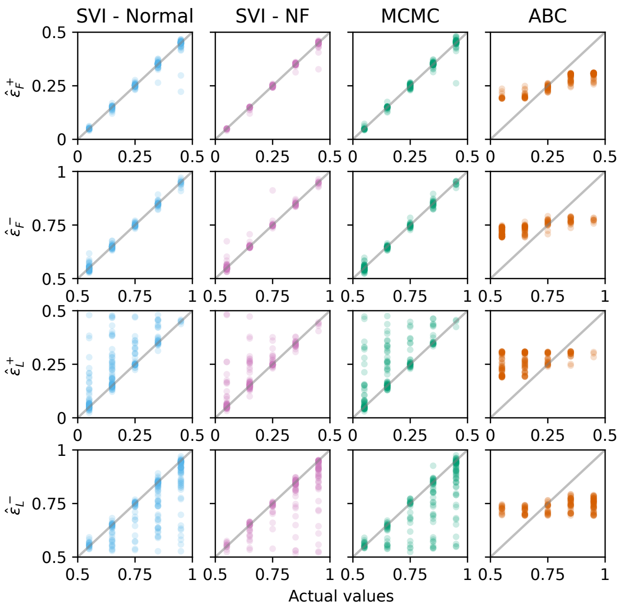

Figure 3: Actual values ε versus estimates $\hat{\varepsilon}$ for each macroscopic parameter (rows) and method (columns).

Parameter Estimates. Figure 3 compares the estimates of the *macroscopic* parameters with their actual values for all the experiments. Summarizing the results, we obtain an average root-mean-square error (RMSE) of 0.044 for SVI with normal, of 0.036 for SVI with NFs, 0.051 for MCMC and 0.125 for ABC. The main hurdle in the inference is discriminating between leaders and followers, both because of the high dimensionality of the parameter and its categorical nature. Since most of the users are followers, the leader's parameters are the hardest to estimate: when the methods are not able to identify the leaders, they tend to set $\hat{\varepsilon}_L^+ = \hat{\varepsilon}_F^+$ and $\hat{\varepsilon}_L^- = \hat{\varepsilon}_F^-$. Since $\varepsilon_F^+ \geq \varepsilon_L^+$, several observations in Figure 3 (row 3) are above the diagonal. Analogously, since $\varepsilon_F^- \leq \varepsilon_L^-$, most of the estimates in Figure 3 (row 4) are below the diagonal.

Estimate Robustness. Figure 4 shows the average error for the four macroscopic parameters of the ABM, and the error rate for the microscopic ones (the roles), i.e., the proportion of roles that are not correctly estimated, as functions of the three hyperparameters that we vary in the experiments. The main insights from Figure 4 are the following:

- SVI with NFs, followed by SVI with normal, outperforms the other competitors in each scenario.
- The three methods based on PGMs (the two SVI methods and MCMC) show better performances compared to the simulation-based method, ABC. In particular, ABC is not able to guess the roles of the agents, with error rates on

---

[1] The code to reproduce these results is available at https://github.com/jaclenti/VI_opinion_dynamics

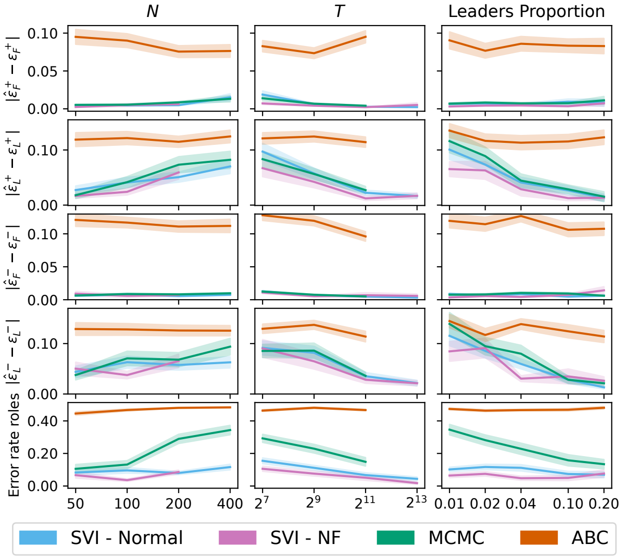

Figure 4: Specific errors of ε_F^+, ε_L^+, ε_F^-, ε_L^-, and r as functions of N (left), T (center), and proportion of leaders (right). The error bars represent the standard errors.

the role microparameters above 40%.

- As T grows, we have more data, thus the estimates of SVI and MCMC improve.
- The error rate of the roles, and consequently all the other errors, grows as N increases. Indeed, as N grows, we have both an increased number of target parameters and a reduced number of observations per agent.
- The VI methods are the only ones able to estimate the role of microparameters with high accuracy. The average error rates for SVI with normal, SVI with NFs, MCMC, and ABC, are 0.09, 0.060, 0.22, and 0.47, respectively. The proportion of experiments with 100% accuracy is, respectively, 0.38, 0.54, 0.29, and 0.

Running time. The estimation times vary considerably among the tested methods (Figure 5). First, both the VI methods scale well with the length of the ABMs trajectories (T). The scalability of VI arises from its optimization-based approach, which avoids the need to explicitly evaluate the entire dataset during inference. This computational advantage becomes more evident as the dataset size grows (Blei, Kucukelbir, and McAuliffe 2017). Second, the estimation time of SVI with NFs grows linearly with N. This reflects the increased number of parameters underlying the neural networks implementing the NFs. Conversely, the estimation time of MCMC increases linearly with T, reflecting the number of samples required to replicate larger datasets. Overall, SVI with normal is the most efficient method: it reduces the estimation time by a factor of 56.0 compared to SVI with NFs, 36.9 compared to MCMC, and 36.5 compared to ABC. On average, the estimation time of SVI with normal in the largest experiments ($N = 400$, $T = 8192$) is 61 seconds. The experiments having SVI with NFs with $N = 400$ and the ones with ABC and MCMC with $T = 8192$ ran over the time limit (3 hours).

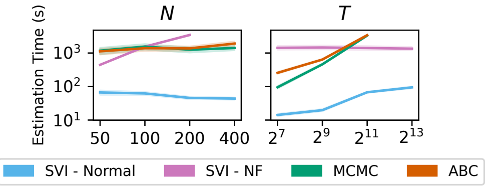

Figure 5: Estimation time as a function of N (left) and T (right). The error bars represent standard errors.

Discussion

Although opinion dynamics models are widely used to investigate online social systems, their application remain constrained by the lack of reliable validation and estimation strategies when dealing with real-world data. A key challenge in this field is the parameter estimation from observed data trajectories. This paper tackles this well-known problem, by presenting a methodology that is able to estimate with high accuracy both the 4-dimensional vector of macro-parameters and the 200 microscopic parameters of the agents' attributes within a Bounded Confidence Model with backfire effect and leaders. Our approach starts from the translation of the ABM into its probabilistic generative counterpart, which we refer to as PGABM (Probabilistic Generative ABM). Since existing inference methods for PGABMs rely on maximum likelihood estimation (Lenti, Monti, and De Francisci Morales 2024) or expectation maximization (Monti, De Francisci Morales, and Bonchi 2020; Monti et al. 2022), their applicability is limited. This is due to the complex non-trivial analytical derivation and maximization of the likelihood. This work marks a significant advancement in the estimation of the ABMs, by showing the use of a complete pipeline of variational inference on the translated ABM. This novelty offers a two-fold advantage compared to previous works on PGABMs. On the one hand, the introduction of a Gumbel-Softmax approximation of the agent states circumvents the non-differentiability of categorical variables. On the other hand, the implementation of Stochastic Variational Inference in a probabilistic programming framework (in our case, NumPyro) allows for the definition of a probabilistic generative model without the necessity of an explicit derivation of the likelihood.

Although our study focuses only on one specific model, we underline that the latter is not only popular but also challenging because of the latent parameters' high dimensionality and their categorical nature. Indeed, it would be unfeasible to visit all the 2^{200} combinations of roles within a simulation-based or likelihood-based approach. Additionally, the presented pipeline required only minimal customization in the translation from the ABM to the PGABM, its implementation in NumPyro, and the optimization of the ELBO. Applying it to other models would be easy, given its flexibility and generality. Finally, the capability to estimate the model parameters with synthetic data paves the way for future works that could investigate opinion dynamics in real-world scenarios.

References

Blei, D. M.; Kucukelbir, A.; and McAuliffe, J. D. 2017. Variational inference: A review for statisticians. *Journal of the American statistical Association*, 112(518): 859–877.

Carrella, E. 2021. No free lunch when estimating simulation parameters. *Journal of Artificial Societies and Social Simulation*, 24(2).

Chen, S.; Xiao, L.; and Kumar, A. 2023. Spread of misinformation on social media: What contributes to it and how to combat it. *Computers in Human Behavior*, 141: 107643.

Cinelli, M.; Morales, G. D. F.; Galeazzi, A.; Quattrociocchi, W.; and Starnini, M. 2020. Echo chambers on social media: A comparative analysis. *arXiv preprint arXiv:2004.09603*.

Cinelli, M.; Pelicon, A.; Mozetič, I.; Quattrociocchi, W.; Novak, P. K.; and Zollo, F. 2021. Dynamics of online hate and misinformation. *Scientific reports*, 11(1): 22083.

Csilléry, K.; Blum, M. G.; Gaggiotti, O. E.; and François, O. 2010. Approximate Bayesian computation (ABC) in practice. *Trends in ecology & evolution*, 25(7): 410–418.

Fagiolo, G.; Guerini, M.; Lamperti, F.; Moneta, A.; and Roventini, A. 2019. Validation of agent-based models in economics and finance. In *Computer simulation validation*, 763–787. Springer.

FORCE11. 2020. The FAIR Data principles. https://force11.org/info/the-fair-data-principles/.

Gebru, T.; Morgenstern, J.; Vecchione, B.; Vaughan, J. W.; Wallach, H.; Iii, H. D.; and Crawford, K. 2021. Datasheets for datasets. *Communications of the ACM*, 64(12): 86–92.

Grazzini, J.; Richiardi, M. G.; and Tsionas, M. 2017. Bayesian estimation of agent-based models. *Journal of Economic Dynamics and Control*, 77: 26–47.

Hoffman, M. D.; Blei, D. M.; Wang, C.; and Paisley, J. 2013. Stochastic variational inference. *Journal of Machine Learning Research*.

Hoffman, M. D.; Gelman, A.; et al. 2014. The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *J. Mach. Learn. Res.*, 15(1): 1593–1623.

Jager, W.; and Amblard, F. 2005. Uniformity, bipolarization and pluriformity captured as generic stylized behavior with an agent-based simulation model of attitude change. *Computational & Mathematical Organization Theory*, 10(4): 295–303.

Jang, E.; Gu, S.; and Poole, B. 2016. Categorical Reparameterization with Gumbel-Softmax. In *International Conference on Learning Representations*.

Kingma, D. P.; and Ba, J. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Lenti, J.; Monti, C.; and De Francisci Morales, G. 2024. Likelihood-Based Methods Improve Parameter Estimation in Opinion Dynamics Models. *Proceedings of the 17th ACM International Conference on Web Search and Data Mining*.

Lux, T.; and Zwinkels, R. C. 2018. Empirical validation of agent-based models. In *Handbook of computational economics*, volume 4, 437–488. Elsevier.

Maddison, C. J.; Mnih, A.; and Teh, Y. W. 2016. The Concrete Distribution: A Continuous Relaxation of Discrete Random Variables. In *International Conference on Learning Representations*.

Monti, C.; De Francisci Morales, G.; and Bonchi, F. 2020. Learning opinion dynamics from social traces. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 764–773.

Monti, C.; Pangallo, M.; Morales, G. D. F.; and Bonchi, F. 2022. On learning agent-based models from data. *arXiv preprint arXiv:2205.05052*.

Papamakarios, G.; Nalisnick, E.; Rezende, D. J.; Mohamed, S.; and Lakshminarayanan, B. 2021. Normalizing flows for probabilistic modeling and inference. *The Journal of Machine Learning Research*, 22(1): 2617–2680.

Peralta, A. F.; Kertész, J.; and Iñiguez, G. 2022. Opinion dynamics in social networks: From models to data. *arXiv preprint arXiv:2201.01322*.

Platt, D. 2020. A comparison of economic agent-based model calibration methods. *Journal of Economic Dynamics and Control*, 113: 103859.

Prieto Curiel, R.; and González Ramírez, H. 2021. Vaccination strategies against COVID-19 and the diffusion of anti-vaccination views. *Scientific Reports*, 11(1): 6626.

Rezende, D.; and Mohamed, S. 2015. Variational inference with normalizing flows. In *International conference on machine learning*, 1530–1538. PMLR.

Sobkowicz, P.; and Sobkowicz, A. 2021. Agent based model of anti-vaccination movements: Simulations and comparison with empirical data. *Vaccines*, 9(8): 809.

Srikrishnan, V.; and Keller, K. 2021. Small increases in agent-based model complexity can result in large increases in required calibration data. *Environmental Modelling & Software*, 138: 104978.

Weisbuch, G.; Deffuant, G.; Amblard, F.; and Nadal, J.-P. 2002. Meet, discuss, and segregate! *Complexity*, 7(3): 55–63.

Paper Checklist

1. For most authors...
   (a) Would answering this research question advance science without violating social contracts, such as violating privacy norms, perpetuating unfair profiling, exacerbating the socio-economic divide, or implying disrespect to societies or cultures? No, this is a theoretical study.
   (b) Do your main claims in the abstract and introduction accurately reflect the paper's contributions and scope? Yes, the abstract and introduction provide the background and the main contributions of the paper.
   (c) Do you clarify how the proposed methodological approach is appropriate for the claims made? Yes.
   (d) Do you clarify what are possible artifacts in the data used, given population-specific distributions? Yes
   (e) Did you describe the limitations of your work? Yes, we describe the limitations in the Discussion.
   (f) Did you discuss any potential negative societal impacts of your work? NA
   (g) Did you discuss any potential misuse of your work? NA
   (h) Did you describe steps taken to prevent or mitigate potential negative outcomes of the research, such as data and model documentation, data anonymization, responsible release, access control, and the reproducibility of findings? We do not use real data.
   (i) Have you read the ethics review guidelines and ensured that your paper conforms to them? Yes.
2. Additionally, if your study involves hypotheses testing...
   (a) Did you clearly state the assumptions underlying all theoretical results? NA
   (b) Have you provided justifications for all theoretical results? Yes.
   (c) Did you discuss competing hypotheses or theories that might challenge or complement your theoretical results? Yes, we compare with the main competitors.
   (d) Have you considered alternative mechanisms or explanations that might account for the same outcomes observed in your study? Yes, we consider the most popular alternatives in the literature.
   (e) Did you address potential biases or limitations in your theoretical framework? Yes, we address the limitations in the Discussion.
   (f) Have you related your theoretical results to the existing literature in social science? The Introduction provides the necessary background from social science.
   (g) Did you discuss the implications of your theoretical results for policy, practice, or further research in the social science domain? Yes, the Discussion describes how our method can be used to link theoretical models with real-world data.
3. Additionally, if you are including theoretical proofs...
   (a) Did you state the full set of assumptions of all theoretical results? NA
   (b) Did you include complete proofs of all theoretical results? NA
4. Additionally, if you ran machine learning experiments...
   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? Yes, we provide the Github link.
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? We provide the most important hyperparameters, without providing all the details.
   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? Yes.
   (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? We provide the total experimental time, in comparison with other techniques.
   (e) Do you justify how the proposed evaluation is sufficient and appropriate to the claims made? Yes.
   (f) Do you discuss what is "the cost" of misclassification and fault (in)tolerance? We explain which experiments are more likely to be misclassified.
5. Additionally, if you are using existing assets (e.g., code, data, models) or curating/releasing new assets, **without compromising anonymity**...
   (a) If your work uses existing assets, did you cite the creators? NA
   (b) Did you mention the license of the assets? NA
   (c) Did you include any new assets in the supplemental material or as a URL? NA
   (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? NA
   (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? NA
   (f) If you are curating or releasing new datasets, did you discuss how you intend to make your datasets FAIR (see FORCE11 (2020))? NA
   (g) If you are curating or releasing new datasets, did you create a Datasheet for the Dataset (see Gebru et al. (2021))? NA
6. Additionally, if you used crowdsourcing or conducted research with human subjects, **without compromising anonymity**...
   (a) Did you include the full text of instructions given to participants and screenshots? NA
   (b) Did you describe any potential participant risks, with mentions of Institutional Review Board (IRB) approvals? NA
   (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? NA
   (d) Did you discuss how data is stored, shared, and deidentified? NA