RETHINKING SOCIAL SIMULATION IN OPEN-POPULATION SYSTEMS: FROM PERSISTENT AGENTS TO COMMUNICATION PROCESSES

Anonymous authors
Paper under double-blind review

ABSTRACT

Large language models (LLMs) have enabled social simulation with persistent agents, where fixed computational individuals maintain internal states and interact over time. However, large-scale social networks is inherently **open-population**: the participating population changes over time, users enter and leave dynamically, and only a sparse subset becomes observable through communication at any given time. In such systems, persistent account identities do not provide continuously observable individual-state trajectories, making recursively maintained psychological states insufficiently grounded in empirical observations. We therefore shift the simulation target from persistent individual trajectories to the observable evolution of social communication, modeling who participates, what they communicate, and how these communications shape subsequent system dynamics. We instantiate this formulation with participant activation from a global user pool, collective state transition, and an LLM policy for communication generation, while explicitly separating environmental inputs from endogenous system evolution. Across five Qwen backbones, our framework reduces average KL divergence by **71.2%** over Direct LLM and **35.0%** over MF-LLM, with boundary and counterfactual experiments further validating the proposed modeling boundaries. Our results support communication-centered simulation as a principled formulation for long-horizon open-population social systems.

1 INTRODUCTION

Large-scale online platforms such as Twitter, Weibo, and live-streaming communities have become major environments for public opinion and collective response (Squazzoni et al., 2014; Freeman, 2004). Large language models (LLMs) increasingly enable simulation of these systems through agents that generate socially situated actions and content (Gao et al., 2024; 2023; Song et al., 2026). A common formulation represents people as persistent agents and carries the social process forward through their individual histories or latent states (Park et al., 2023; 2024; Piao et al., 2025). We call this a *human-trajectory-centered* formulation. Its empirical interpretation, however, remains unclear (Argyle et al., 2023; Bisbee et al., 2024). *If an LLM agent is intended to stand for a person, the validity of its trajectory appears to depend on an unverified equivalence between the language model and the human participant*. If it is not, the empirical referent of the generated individual trajectory becomes difficult to specify. The central question is therefore whether persistent individual trajectories are the appropriate recursive object for an open social process.

This question is particularly pressing on open online platforms. For any given social event, the potential participant population is massive, while actual participation is sparse and time-varying (Jin et al., 2013; Lu et al., 2018b). Users may enter at any time and contribute only briefly before disengaging, even while remaining active elsewhere on the platform (Szell et al., 2014). Individuals are therefore observed intermittently within an event, whereas the platform process continues through posts, replies, and shares (Kwak et al., 2010; Notarmuzi et al., 2022). These *communicative events* accumulate into a system-visible history that conditions subsequent activity. Although a human-trajectory-centered formulation can be coherent for relatively closed populations with sustained participation, its interpretation becomes less clear in this open setting. It leaves three basic

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

scientific questions unresolved: *what the simulator predicts*, *what it is responsible for*, and *what the LLM represents* (Grimm et al., 2005; Edmonds & Hales, 2003).

What the simulator predicts. A human-trajectory-centered formulation requires individual trajectories to be continued over time, yet in open online systems individuals are only observed when they act. The unobserved evolution between actions therefore has neither direct supervision nor an observable target for validation. This leaves the simulator with a dilemma: either generate latent individual evolution that cannot be empirically verified, or obtain future participants and behavior from the realized trajectory, thereby revealing part of the future.

What the simulator is responsible for. A human-trajectory-centered formulation treats future social change as the continued evolution of its modeled individuals. In an open system, however, future trajectories are shaped by contingencies that have not yet entered the social process. Absorbing these contingencies into individual trajectories collapses possible futures into the single realized continuation the simulator is expected to reproduce. It therefore becomes unclear whether the model is simulating possible social evolution or reconstructing the future that happened to occur.

What the LLM represents. A human-trajectory-centered formulation requires some model to carry each individual's trajectory forward, and LLM-based simulators assign this role to the LLM. Yet the simulator typically observes only sparse digital traces of a person, such as profile attributes and past activity. These observations may support conditional generation of observable communicative acts, but they do not establish that the LLM recovers the person's psychological evolution. The resulting trajectory may serve as a computational state, but the available observations do not justify interpreting it as the recovered psychological evolution of the represented person.

These are not independent limitations. They arise from the same modeling decision: asking an unobserved individual trajectory to carry the continuity of an observable event history. We therefore shift the recursive object of simulation. *Persons remain the participants through whom communicative events are realized, but their latent psychological trajectories are not treated as persistent system state*. Instead, given a system-visible history and declared environmental inputs, the simulator recursively updates that history and estimates a conditional distribution over subsequent observable record updates. Its inputs are observed or declared, its outputs are empirically addressable, and it is responsible for the generated platform process conditional on environmental inputs, not for inventing external events. The empirical claim therefore no longer depends on establishing equivalence between an LLM and a human mind. We arrive at this formulation from the computational requirements of open-population simulation. Conceptually, however, its central distinction echoes Luhmann's social systems theory: social processes reproduce through communication rather than through the continuous evolution of individual consciousness (Luhmann, 1995; Esposito, 2022).

Our architecture realizes this contract through three components. We model **endogenous participation** over an open population, learn an **autonomous communication policy** that generates actions and content from information available at simulation time, and introduce **future-constrained communication learning** to connect locally plausible acts with subsequent process evolution. Future dynamics are used only as training supervision; at inference, the simulator generates and selects candidates without access to the realized future trajectory. These components recursively generate the open platform process rather than reconstructing presumed human trajectories.

Our contributions are threefold.

- We identify a structural mismatch between human-trajectory-centered simulation and open online social systems: ambiguity about prediction, endogenous responsibility, and the interpretation of LLM agents arises from using unobserved individual trajectories to carry a process observed through intermittent platform events.
- We reformulate open online social simulation as recursive prediction and update of a system-visible event history, with declared conditioning information, event-level prediction targets, and an operational system-environment boundary, without requiring equivalence between an LLM and a human mind, in structural alignment with Luhmann's **social systems theory**.
- We instantiate and evaluate this formulation with an open-population simulator combining endogenous participant realization, autonomous action-and-content generation, and future-constrained learning. Across five Qwen backbones, the resulting framework improves long-horizon trajectory fidelity without access to the realized future participant sequence.

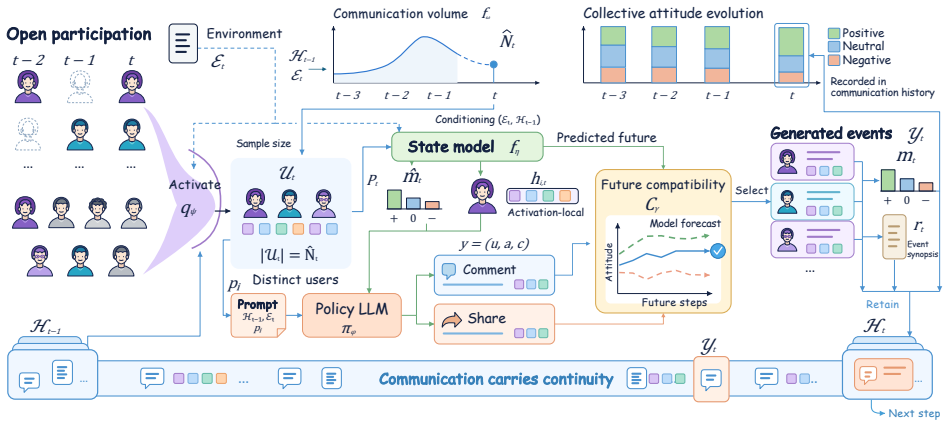

Figure 1: **Overview of OpenSim.** OpenSim recursively models participation, mean-field state transition, and future-constrained communication generation, with realized communication updating the collective state and platform history for subsequent steps.

2 Background and Problem Formulation

We formalize open-population social simulation with sparse, time-varying participation (§2.1), examine limitations of persistent-agent dynamics in LLM-based simulators (§2.2), and define a simulation target for observable communication with an explicit system–environment boundary (§2.3).

2.1 Open-Population Social Systems

We consider an online social process evolving over discrete time steps $t \in \{1, \dots, T\}$. Let $\mathcal{U}$ denote a global pool of potential participants.Unlike closed-agent simulations, the users participating in an online event change over time. At each step, only a sparse and time-varying subset $\mathcal{U}_t \subseteq \mathcal{U}$ becomes active, with $N_t = |\mathcal{U}_t|$ (Lu et al., 2018a; Nie et al., 2018). The same user may reappear across non-consecutive steps. Let $\mathcal{H}_{t-1}$ denote the observable platform history before step t, and $\mathcal{E}_t$ the exogenous information introduced at that step. For each active user $u_i \in \mathcal{U}_t$, a communicative event is represented as $y_{i,t} = (u_i, a_{i,t}, c_{i,t})$, where $a_{i,t}$ is the platform action and $c_{i,t}$ is the expressed or propagated content. The realized event set is $\mathcal{Y}_t = \{y_{i,t} : u_i \in \mathcal{U}_t\}$, which is incorporated into the observable platform history $\mathcal{H}_t$.

We call this an **open-population social system**: participation is sparse and intermittent, with a changing active population. Persistent account identifiers do not imply continuously observable individual-state trajectories; continuity instead resides in the observable platform history $\mathcal{H}_t$.

2.2 Persistent-Agent Formulation

Existing LLM-based social simulators commonly assume a *closed* agent population, where a fixed set of computational agents is instantiated before the rollout and persistently maintained throughout the simulation (Zhang et al., 2025a; Yang et al., 2025). Each agent i maintains an internal state $z_{i,t}$, which may encode opinion, emotion, memory, or other latent attributes, and evolves recursively as

$$z_{i,t+1} \sim G_\theta(z \mid z_{i,t}, p_i, \mathcal{H}_t), \qquad y_{i,t+1} \sim \pi_\phi(y \mid z_{i,t+1}, p_i, \mathcal{H}_t), \tag{1}$$

where p_i denotes system-visible participant information. Such information may condition observable communication, but does not by itself identify a private psychological trajectory. The resulting process is carried forward as $\{z_{i,t}\} \to \{z_{i,t+1}\} \to \mathcal{Y}_{t+1} \to \mathcal{H}_{t+1}$.

This formulation differs fundamentally from the open-population setting, where only a time-varying subset is active at each step. For a user absent between two active steps, the system provides no direct observation of the intermediate individual state, so recursively carrying forward $z_{i,t}$ propagates simulator-generated latent states beyond what is identified by the observable platform history. Fixing the same agents throughout the rollout reduces the problem to a closed-population approximation,

whereas supplying future active users from the realized trajectory introduces future participation information (Mi et al., 2025; Zhang et al., 2026a). Thus, in an open-population system, **both *who participates* and *what they communicate* must be generated as part of the social process**.

2.3 Simulation Target and Boundary

We therefore shift the recursive object from persistent individual trajectories to the observable evolution of communication. Given the platform history $\mathcal{H}_t$ and declared environmental input $\mathcal{E}_{t+1}$, the simulator generates

$$\mathcal{Y}_{t+1} \sim p_\theta(\cdot \mid \mathcal{H}_t, \mathcal{E}_{t+1}), \qquad \mathcal{H}_{t+1} = F(\mathcal{H}_t, \mathcal{Y}_{t+1}, \mathcal{E}_{t+1}). \tag{2}$$

Accordingly, the recursive process becomes $\mathcal{H}_t \rightarrow \mathcal{Y}_{t+1} \rightarrow \mathcal{H}_{t+1}$, rather than a recursion over persistent latent individual states.

This formulation separates latent individual state $z_{i,t}$, participant activation $\mathcal{U}_t$, and environmental input $\mathcal{E}_t$. Latent individual states are not treated as persistent recursive system state; participant activation is generated as part of the simulated process; and environmental inputs may affect the process without being endogenous prediction targets. Operationally, variables generated by the simulator and recursively incorporated into $\mathcal{H}$ belong to the simulated system, whereas externally supplied variables whose generation is not modeled belong to $\mathcal{E}$.

Although motivated by the computational requirements of open-population simulation, this formulation structurally corresponds to Luhmann's view that social systems reproduce through communication rather than through the continuous evolution of individual consciousness (Luhmann, 1995). From this perspective, the continuity of a social process can be located in recursively connected communications, rather than in continuously maintained internal states of persistent individuals. Esposito's notion of *artificial communication* further suggests that algorithmic systems may participate in producing communication without being treated as equivalents of human consciousness (Esposito, 2022). This distinction is particularly relevant to LLM-based simulation: an LLM may model the production of observable communication without requiring the stronger assumption that it constitutes a persistent psychological surrogate for a real participant.

3 Methodology

We instantiate the proposed formulation with three components: participation modeling (§3.1), collective state transition (§3.2), and communicative policy modeling (§3.3). These components jointly govern participation, collective state, and communication at each simulation step, as summarized in Algorithm 1 in the appendix.

3.1 Participation Modeling

The first component models the scale and composition of communication at time t by forecasting the communication volume and activating the participants who realize these events.

Communication Volume Forecasting Let N_t denote the number of communicative events at time t. We forecast the communication volume using information available before the current step:

$$\hat{N}_t = f_\omega(\mathcal{H}_{t-1}, \mathcal{E}_t), \tag{3}$$

where $\mathcal{H}_{t-1}$ is the accumulated platform history and $\mathcal{E}_t$ is the declared environmental input. The forecasting model summarizes recent communication dynamics and event-level context without accessing the realized future participation volume.

Global-Pool Participant Activation Given $\hat{N}_t$, participants are activated from a predefined global user pool $\mathcal{U} = \{u_1, u_2, \ldots, u_M\}$. Each candidate u_i is represented by system-visible information p_i, such as profile attributes, interest labels, and historical activity. We assign each candidate an activation likelihood

$$q_\psi(u_i \mid p_i, \mathcal{H}_{t-1}, \mathcal{E}_t), \qquad u_i \in \mathcal{U}, \tag{4}$$

and obtain $\mathcal{U}_t$ by sampling $\hat{N}_t$ participants without replacement according to these likelihoods, such that $|\mathcal{U}_t| = \hat{N}_t$. Each activated participant realizes one communicative event at the current step.

Activation is step-specific, allowing users to appear once, reappear after inactivity, or never reappear; process-level continuity is carried by $\mathcal{H}_t$ rather than by persistently maintained agents.

3.2 Mean-Field State Transition

Following M^{2}Sim (Zhang et al., 2026a), we represent the collective communication state with a distributional mean field $m_{t-1} \in \Delta^2$ and a textual synopsis r_{t-1} retained in $\mathcal{H}_{t-1}$. Unlike persistent-agent formulations, participant-level latent representations are inferred only for users activated at the current step.

Given the current participant features P_t, platform history $\mathcal{H}_{t-1}$, and environmental input $\mathcal{E}_t$, the transition model predicts

$$(\hat{m}_t, \{h_{i,t}\}_{u_i \in \mathcal{U}_t}) = f_\eta\left(\mathcal{H}_{t-1}, \mathcal{E}_t, P_t\right), \tag{5}$$

where $\hat{m}_t$ is the predicted attitude composition of the current communication and $h_{i,t}$ is an activation-local latent representation used by the communication policy. These representations are not recursively maintained across inactive periods.

After the communicative events $\mathcal{Y}_t$ are generated, their expressed attitudes are aggregated to update the realized mean field m_t, while the textual synopsis is updated as

$$r_t = g_\xi\left(r_{t-1}, m_t, \mathcal{Y}_t, \mathcal{E}_t\right). \tag{6}$$

Thus, $\hat{m}_t$ guides current communication, whereas (m_t, r_t) records the realized simulated process and is incorporated into $\mathcal{H}_t$ for subsequent steps.

3.3 Future-Constrained Communication Policy

Given an activated participant $u_{t,j}$, the policy generates an action and its associated content:

$$(a_{t,j}, c_{t,j}) \sim \pi_\phi\big(\cdot \mid \mathcal{H}_{t-1}, \mathcal{E}_t, p_{u_{t,j}}, h_{t,j}, \hat{m}_t\big), \tag{7}$$

yielding the communicative event $y_{t,j} = (u_{t,j}, a_{t,j}, c_{t,j})$.

Local likelihood alone does not ensure that a plausible response is compatible with subsequent collective evolution. Building on the long-horizon consistency objective of M^{2}Sim (Zhang et al., 2026a), we introduce a response-sensitive trajectory critic

$$\widehat{\mathbf{m}}_{t+1:t+K} = C_\gamma\left(\mathcal{H}_{t-1}, \hat{m}_t, y_{t,j}\right), \tag{8}$$

which predicts the subsequent mean-field trajectory conditioned on the candidate communication. Given the observed training trajectory $\mathbf{m}_{t+1:t+K}$, the trajectory loss is

$$\mathcal{L}_{\text{traj}} = \sum_{k=1}^{K} D(m_{t+k}, \widehat{m}_{t+k}), \tag{9}$$

where $D(\cdot,\cdot)$ measures distributional discrepancy.

A trajectory critic may otherwise rely on the shared context while ignoring the response itself. We therefore construct masked and cross-context mismatched responses $\tilde{y}_{t,j}$ and encourage the observed response to better match its associated future trajectory:

$$\mathcal{L}_{\text{sens}} = \max\big(0,\, \delta + D(\mathbf{m}_{t+1:t+K}, C_\gamma\left(\mathcal{H}_{t-1}, \hat{m}_t, y_{t,j}\right)) - D(\mathbf{m}_{t+1:t+K}, C_\gamma\left(\mathcal{H}_{t-1}, \hat{m}_t, \tilde{y}_{t,j}\right))\big). \tag{10}$$

After communication SFT and critic initialization, the policy adapters and trajectory critic are jointly optimized with

$$\mathcal{L}_{\text{joint}} = \mathcal{L}_{\text{comm}} + \lambda_{\text{traj}}\mathcal{L}_{\text{traj}} + \lambda_{\text{sens}}\mathcal{L}_{\text{sens}}. \tag{11}$$

At inference time, candidate communications are ranked by the critic according to their compatibility with the state model's predicted future mean-field evolution. The selected events constitute $\mathcal{Y}_t$ and are used to update the realized mean field and textual synopsis.

Table 1: **Main results across backbone models.** OpenSim / OpenSim† use event-specific / external participant pools. *: observed future participant sequence. **Bold**/underline: best/second-best per model. Parentheses: improvement (%) over gray-shaded MF-LLM; green/red = better/worse. Darker purple = larger gains.

Model	Method	KL ↓	Wass. ↓	DTW ↓	Macro-F1 ↑	Micro-F1 ↑	NLL ↓
Qwen 2-1.5B	Direct LLM*	2.4090	0.3336	0.3055	0.3441	0.5686	4.6838
	Social Retrieval*	1.6678	0.2568	0.2063	0.3636	0.5526	5.1355
	MF-LLM*	1.2494	0.2246	0.1773	0.4311	0.6295	4.6060
	OpenSim†	2.0773 (-66.3%)	0.2905 (-29.3%)	0.2256 (-27.2%)	0.4209 (-2.4%)	0.6199 (-1.5%)	**4.1414** (+10.1%)
	OpenSim	**1.1352** (+9.1%)	**0.1922** (+14.4%)	**0.1261** (+28.9%)	**0.4444** (+3.1%)	**0.6616** (+5.1%)	4.3377 (+5.8%)
Qwen 2.5-1.5B	Direct LLM*	3.3485	0.2844	0.3384	0.3730	0.4651	6.3707
	Social Retrieval*	2.4112	0.2243	0.2695	0.4569	0.5566	6.2568
	MF-LLM*	1.4739	0.1941	0.1396	0.5408	0.5881	5.1429
	OpenSim†	1.5739 (-6.8%)	0.2142 (-10.4%)	0.1934 (-38.5%)	0.5001 (-7.5%)	0.6008 (+2.2%)	**4.4138** (+14.2%)
	OpenSim	**1.0667** (+27.6%)	**0.1724** (+11.2%)	**0.1252** (+10.3%)	**0.6097** (+12.7%)	**0.6252** (+6.3%)	4.5980 (+10.6%)
Qwen 2.5-3B	Direct LLM*	3.8755	0.3603	0.3771	0.3138	0.4561	13.9996
	Social Retrieval*	1.8466	0.2283	0.2591	0.3817	0.5129	14.0303
	MF-LLM*	1.5061	0.1546	0.2311	0.4598	0.6222	4.5446
	OpenSim†	1.5030 (+0.2%)	0.1495 (+3.3%)	0.1824 (+21.1%)	0.4652 (+1.2%)	0.6545 (+5.2%)	**4.4980** (+1.0%)
	OpenSim	**1.0048** (+33.3%)	**0.1273** (+17.7%)	**0.1369** (+40.8%)	**0.5058** (+10.0%)	**0.7278** (+17.0%)	4.5356 (+0.2%)
Qwen 2.5-7B	Direct LLM*	3.0914	0.3107	0.3097	0.3471	0.5581	13.0946
	Social Retrieval*	2.5142	0.2981	0.2882	0.3701	0.5821	12.8823
	MF-LLM*	1.6049	0.2234	0.2564	0.4380	0.6117	**4.6823**
	OpenSim†	1.4294 (+10.9%)	0.2289 (-2.5%)	0.2509 (+2.1%)	0.4455 (+1.7%)	0.6340 (+3.6%)	4.9849 (-6.5%)
	OpenSim	**0.8509** (+47.0%)	**0.1759** (+21.3%)	**0.1572** (+38.7%)	**0.4855** (+10.8%)	**0.6752** (+10.4%)	5.0921 (-8.8%)
Qwen 3-1.7B	Direct LLM*	5.1480	0.4279	0.4191	0.3314	0.5236	25.7470
	Social Retrieval*	4.7743	0.3453	0.2210	0.3478	0.5451	25.4819
	MF-LLM*	2.0902	0.2549	0.1715	0.5095	0.6518	6.3621
	OpenSim†	1.3115 (+37.3%)	0.2377 (+6.7%)	0.1583 (+7.7%)	0.4865 (-4.5%)	0.6825 (+4.7%)	**4.4081** (+30.7%)
	OpenSim	**1.0929** (+47.7%)	**0.1780** (+30.2%)	**0.1098** (+36.0%)	**0.5445** (+6.9%)	**0.7610** (+16.8%)	4.4414 (+30.2%)

4 EXPERIMENTS

We evaluate the proposed framework from four perspectives: (1) long-horizon simulation against representative LLM-based social simulation baselines across multiple backbones (§4.2); (2) environmental communication interventions and their effects on subsequent system evolution (§4.3); (3) boundary analysis through perturbations of participant priors and environmental inputs (§4.4); and (4) ablations of key components for simulation fidelity and long-horizon consistency (§4.5).

4.1 EXPERIMENTAL SETUP

Dataset. We use an event-level Chinese social-media corpus with intervention-aware annotations, macro-level interaction volume and attitude trajectories, and fine-grained micro-level user interactions. This structure supports joint modeling of external interventions, collective dynamics, and individual communication behavior. Details are provided in Appendix D.1.

Backbone Models. We evaluate five backbone settings to examine robustness across model families and scales: Qwen2-1.5B, Qwen2.5-1.5B, Qwen2.5-3B, Qwen2.5-7B, and Qwen3-1.7B. All methods within the same backbone setting follow the same evaluation protocol.

Metrics. We evaluate both distributional trajectory fidelity and action-level prediction. For collective dynamics, we report KL divergence (**KL**), Wasserstein distance (**Wass.**), and Dynamic Time Warping distance (**DTW**), where lower values indicate better trajectory agreement. For action-level prediction, we report Macro-F1, Micro-F1, and negative log-likelihood (**NLL**).

4.2 DO WE NEED FUTURE PARTICIPANTS?

Baselines. We compare OpenSim with three representative social simulation baselines, *all of which use the observed participant sequence from the realized trajectory, leaking future participation information*: (1) **Direct LLM***, which generates actions directly from the event context (Gao et al., 2023; Zhang et al., 2025b); (2) **Social Retrieval***, which conditions generation on retrieved social traces (Piao et al., 2025; Yang et al., 2025); and (3) **MF-LLM***, which uses a recursively updated textual mean field to guide generation (Mi et al., 2025). For OpenSim, we evaluate two participant

Table 2: **Communication interventions across backbones.** ORIGINAL retains the original intervention, REMOVED removes it, and OPPOSITE replaces it with an intervention of the opposite stance. **Bold**/underline indicate the best/second-best agreement with the corresponding reference.

Model	Intervention	KL	Wass.	DTW	Macro-F1	Micro-F1	NLL	Att. Acc. ↑	Att. DTW ↓
Qwen 2-1.5B	ORIGINAL	**1.1352**	**0.1922**	**0.1261**	**0.4444**	**0.6616**	**4.3377**	**64.4**	**0.1779**
	REMOVED	3.4784	0.4974	0.4528	0.3502	0.5383	4.7429	49.6	0.3235
	OPPOSITE	2.0065	0.4126	0.3652	0.3539	0.5509	4.6957	35.1	0.4242
Qwen 2.5-1.5B	ORIGINAL	**1.0667**	**0.1724**	**0.1252**	**0.6097**	**0.6252**	**4.5980**	**65.6**	**0.1717**
	REMOVED	2.4076	0.3495	0.2864	0.3771	0.5656	5.1793	36.8	0.4212
	OPPOSITE	2.2847	0.3693	0.3089	0.3741	0.5653	5.4473	35.1	0.4671
Qwen 3-1.7B	ORIGINAL	**1.0929**	**0.1780**	**0.1098**	**0.5445**	**0.7610**	**4.4414**	**69.1**	**0.1241**
	REMOVED	3.9696	0.2611	0.2606	0.3804	0.6870	4.7457	41.7	0.4810
	OPPOSITE	2.0284	0.3300	0.3328	0.3570	0.6582	4.7813	36.4	0.4942

Notes. Attitude metrics are computed on 57 common supported event-days after warm-up. Att. Acc. measures dominant-class agreement, and Att. DTW uses path-normalized Euclidean costs. Both are text-derived attitude proxies rather than native state-transition states.

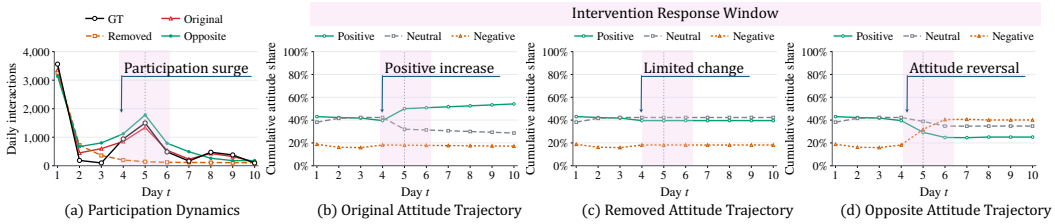

Figure 2: **Example of communication intervention dynamics.** (a) Participation dynamics across intervention settings. (b–d) Cumulative attitude trajectories under ORIGINAL, REMOVED, and OPPOSITE, respectively. The dashed line marks the intervention point.

pools: **OpenSim** uses an event-specific participant pool constructed from the corresponding event's interaction records, while **OpenSim**† uses an external participant pool of 15,992 real-user profiles from SocioVerse (Zhang et al., 2025a). Table 1 reports the main simulation results.

(1) OpenSim remains superior without access to the realized future participant sequence. OpenSim uses only an event-specific candidate pool and generates participation during simulation, whereas all baselines are supplied with the observed participant sequence from the realized trajectory. Despite this information disadvantage, OpenSim consistently achieves substantially better trajectory fidelity, reducing average KL by **71.2%** over Direct LLM and **35.0%** over MF-LLM. This shows that accurate long-horizon simulation does not require knowing in advance which users will participate at each future step.

(2) Exact event-specific participants are not necessary for recovering collective dynamics. Even when OpenSim† uniformly samples participants from an event-agnostic global real-user pool, its performance remains close to MF-LLM, which is given the realized participant sequence. Their average KL is nearly identical (1.5790 vs. 1.5849), while OpenSim† also achieves comparable trajectory distances and action-level performance. This suggests that sparse system-visible user information can be sufficient for reproducing collective dynamics without requiring the simulator to recover the exact individuals who appear in the realized event trajectory.

(3) These patterns are consistent across backbone models. The same conclusions hold across all five Qwen backbones, spanning Qwen2, Qwen2.5, and Qwen3 from 1.5B to 7B parameters. OpenSim consistently outperforms sequence-conditioned baselines in trajectory fidelity, while OpenSim† remains competitive under the substantially weaker global-pool setting. The gains therefore reflect the proposed simulation formulation rather than a particular backbone or model scale.

4.3 HOW DO COMMUNICATION INTERVENTIONS RESHAPE SYSTEM EVOLUTION?

Setting: Controlled Communication Interventions. For each test event, we identify a consequential external communication and construct three conditions while keeping the pre-intervention history unchanged: ORIGINAL retains the intervention, REMOVED removes it, and OPPOSITE replaces

Table 3: **Boundary-controlled comparison across backbones.** DECLARED introduces consequential external information as a declared environmental input, whereas SIMULATED requires the simulated population to generate it. GENERATED predicts participation during rollout, whereas FUTURE SEQUENCE supplies the observed participant sequence from the realized trajectory. † denotes the participant–environment assumptions corresponding to M^{2}Sim (Zhang et al., 2026a). **Bold**/underline indicate the best/second-best result within each backbone.

Model	Environment	Participation	KL ↓	Wass. ↓	DTW ↓	Macro-F1 ↑	Micro-F1 ↑	NLL ↓
Qwen 2-1.5B	DECLARED	GENERATED	1.1352	0.1922	0.1261	0.4444	**0.6616**	4.3377
	DECLARED	FUTURE SEQUENCE	**0.7686**	**0.1869**	**0.1219**	**0.4479**	0.6416	**4.0911**
	SIMULATED	GENERATED	1.2633	0.2045	0.1392	0.4149	0.6059	4.3448
	SIMULATED	FUTURE SEQUENCE†	0.8575	0.1892	0.1240	0.4260	0.6363	4.1696
Qwen 2.5-1.5B	DECLARED	GENERATED	1.0667	0.1724	**0.1252**	0.6097	0.6252	4.5980
	DECLARED	FUTURE SEQUENCE	**0.7395**	**0.1674**	0.1295	**0.6339**	**0.6471**	4.3250
	SIMULATED	GENERATED	1.3847	0.1906	0.1376	0.5977	0.6064	4.3886
	SIMULATED	FUTURE SEQUENCE†	0.9327	0.1860	0.1301	0.6080	0.6128	**4.2038**
Qwen 3-1.7B	DECLARED	GENERATED	1.0929	0.1780	**0.1098**	0.5445	0.7610	4.4414
	DECLARED	FUTURE SEQUENCE	**0.6202**	**0.1726**	0.1168	**0.5772**	**0.7641**	**4.2779**
	SIMULATED	GENERATED	1.3370	0.1989	0.1249	0.5235	0.7339	4.3599
	SIMULATED	FUTURE SEQUENCE†	0.7269	0.1817	0.1240	0.5459	0.7411	4.3711

it with an intervention of the opposite stance. After the intervention point, participant activation, state transition, and communication generation are autonomously rolled out by the same simulator. ORIGINAL uses one rollout per event, while REMOVED and OPPOSITE average four rollouts.

Facet A: Removal primarily attenuates participation and distributional support. REMOVED produces the largest KL divergence across all three backbones, averaging 3.2852, compared with 2.1065 for OPPOSITE and 1.0983 for ORIGINAL. Removing a consequential communication suppresses the downstream activation associated with that intervention, reducing participation and distributional support in the subsequent process. This produces the strongest overall distributional mismatch, while the resulting trajectory changes are comparatively less directional.

Facet B: Opposite interventions redirect attitude and temporal evolution. OPPOSITE preserves the intervention signal but reverses its stance, causing attitudes, emotions, and expressed positions to evolve in the opposite direction. On Qwen2.5-1.5B and Qwen3-1.7B, it produces larger Wasserstein and DTW deviations than REMOVED; it yields the largest attitude discrepancy on every backbone, with the lowest average Att. Acc. (35.5%) and highest Att. DTW (0.4618), compared with 42.7% and 0.4086 for REMOVED. Thus, removal primarily attenuates the communication process, whereas an opposite intervention more strongly redirects its semantic and attitudinal trajectory.

4.4 WHAT INFORMATION SHOULD THE SIMULATOR RECEIVE?

Setting: Boundary-Controlled Comparison. We isolate two modeling boundaries while keeping the state-transition and communication-policy modules fixed: (1) whether participation is generated by the simulator or supplied from the future participant sequence in the realized trajectory, and (2) whether consequential external information enters as a declared environmental input or is reconstructed by a simulated information-source agent. Their Cartesian product yields four controlled settings. Under this decomposition, M^{2}Sim (Zhang et al., 2026a) corresponds to the boundary assumptions that combine simulated information-source generation with the observed future participant sequence. Table 3 reports the results across three backbones.

Fact C: Declared environmental inputs preserve the system–environment boundary. When participation is generated, supplying consequential external information as a declared environmental input reduces average KL from 1.3283 to 1.0983 (17.3%), Wasserstein distance from 0.1980 to 0.1809, and DTW from 0.1339 to 0.1204. The same pattern remains when the future participant sequence is supplied, where average KL decreases from 0.8390 to 0.7094. External contingencies are therefore modeled more faithfully when they enter through the declared environment and the simulator predicts the communication that follows, rather than being reconstructed as endogenous products of the simulated population.

Fact D: Future participation information improves reconstruction accuracy. Supplying the realized future participant sequence substantially improves trajectory matching under both environ-

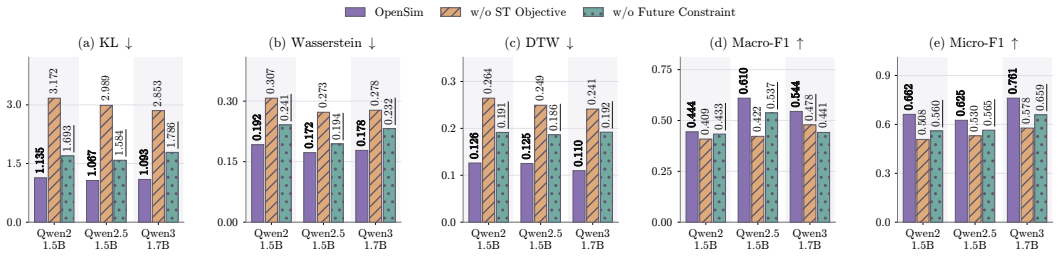

Figure 3: **Component ablations across backbones.** We remove state-transition supervision or the future-constrained policy objective. **Bold**/underlined values mark the best/second-best results per backbone. Full results: Appendix Table 7.

mental settings. With declared environmental inputs, average KL decreases from 1.0983 to 0.7094 (35.4%); when the external information is instead generated by a simulated source, it decreases from 1.3283 to 0.8390 (36.8%). The numerical gain therefore comes from information unavailable in genuine forward simulation: knowing in advance who will participate. This explains why future-conditioned settings can achieve stronger reconstruction accuracy while solving an easier problem than the open-population simulation considered in our main experiments.

4.5 What Drives OpenSim? Component Ablations

Setting: Component Ablations. We ablate two objectives while keeping participation modeling and the remaining simulation pipeline unchanged: w/o ST Objective removes the state-transition objective, while w/o Future Constraint removes the future-constrained policy objective. Figure 3 reports results across three backbones.

Fact E: State-transition supervision anchors long-horizon collective dynamics. Removing the state-transition objective causes the largest degradation across all three backbones. Average KL increases from 1.0983 to 3.0043 (+173.5%), Wasserstein distance from 0.1809 to 0.2858 (+58.0%), and DTW from 0.1204 to 0.2517 (+109.1%). Average Micro-F1 falls from 0.6826 to 0.5385, indicating that the degradation extends across distributional and predictive metrics. These results show that the mean-field state transition provides the primary cross-step structure for maintaining coherent collective evolution, linking successive population states throughout the simulated trajectory.

Fact F: Future constraints improve trajectory fidelity beyond local likelihood. Removing the future constraint increases average KL to 1.6877 (+53.7%) and DTW to 0.1899 (+57.8%), while reducing average Micro-F1 to 0.5943. In contrast, its average NLL improves from 4.4590 to 4.2833, with particularly lower NLL on Qwen2.5-1.5B and Qwen3-1.7B. The divergence between local likelihood and trajectory metrics shows that more likely communications do not necessarily yield more faithful long-horizon social evolution, motivating the future-constrained communication objective.

5 Conclusion

We revisit what should serve as the recursive object of simulation in open online social systems. Rather than carrying the process through persistent synthetic individuals and their latent trajectories, we formulate social simulation around the observable evolution of communication, where participation, collective state, and communicative events are generated from information available at simulation time under an explicit system–environment boundary. This formulation preserves individuals as participants without requiring their continuously evolving psychological states to function as persistent system state. Its structure is also consistent with a communication-centered view of social systems, while remaining a computational formulation derived from the open-population setting itself. We hope this perspective provides a clearer foundation for building and evaluating long-horizon social simulators whose predictions remain empirically addressable and whose modeling responsibilities are explicitly defined.

AI USE STATEMENT

We used generative AI tools to assist with language editing, translation, and improving the clarity of the manuscript. LLMs were also used as components of the simulation framework and for semantic annotation in evaluation, as described in the methodology and experimental settings. All AI-assisted content was reviewed by the authors, who take full responsibility for the final manuscript.

ETHICS STATEMENT

This work studies open-population social simulation for analyzing collective communication and intervention responses. Such techniques could be misused for targeted persuasion or manipulation at scale. Simulated communications and states should not be interpreted as verified psychological attributes of real individuals, and LLM-generated outputs may contain biases or errors. We therefore encourage appropriate ethical review, transparency, and privacy safeguards in real-world applications.

REFERENCES

Lisa P Argyle, Ethan C Busby, Nancy Fulda, Joshua R Gubler, Christopher Rytting, and David Wingate. Out of one, many: Using language models to simulate human samples. *Political Analysis*, 31(3):337–351, 2023.

James Bisbee, Joshua D Clinton, Cassy Dorff, Brenton Kenkel, and Jennifer M Larson. Synthetic replacements for human survey data? the perils of large language models. *Political Analysis*, 32 (4):401–416, 2024.

Riley Crane and Didier Sornette. Robust dynamic classes revealed by measuring the response function of a social system. *Proceedings of the National Academy of Sciences*, 105(41):15649–15653, 2008.

Bruce Edmonds and David Hales. Replication, Replication and Replication: Some Hard Lessons from Model Alignment. *Journal of Artificial Societies and Social Simulation*, 6(4), 2003.

Elena Esposito. *Artificial Communication: How Algorithms Produce Social Intelligence*. The MIT Press, Cambridge, MA, 2022. doi: 10.7551/mitpress/14189.001.0001.

Linton Freeman. The development of social network analysis. *A Study in the Sociology of Science*, 1(687):159–167, 2004.

Chen Gao, Xiaochong Lan, Zhihong Lu, Jinzhu Mao, Jinghua Piao, Huandong Wang, Depeng Jin, and Yong Li. S$\hat{3}$: Social-network Simulation System with Large Language Model-Empowered Agents. *arXiv preprint arXiv:2307.14984*, 2023.

Chen Gao, Xiaochong Lan, Nian Li, Yuan Yuan, Jingtao Ding, Zhilun Zhou, Fengli Xu, and Yong Li. Large language models empowered agent-based modeling and simulation: A survey and perspectives. *Humanities and Social Sciences Communications*, 11(1):1–24, 2024.

Volker Grimm, Eloy Revilla, Uta Berger, Florian Jeltsch, Wolf M Mooij, Steven F Railsback, Hans-Hermann Thulke, Jacob Weiner, Thorsten Wiegand, and Donald L DeAngelis. Pattern-oriented modeling of agent-based complex systems: lessons from ecology. *science*, 310(5750):987–991, 2005.

Alan G Hawkes. Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1):83–90, 1971.

Xinran He, Theodoros Rekatsinas, James Foulds, Lise Getoor, and Yan Liu. Hawkestopic: A joint model for network inference and topic modeling from text-based cascades. In *International conference on machine learning*, pp. 871–880. PMLR, 2015.

Renhong Huang, Ning Tang, Jiarong Xu, Yuxuan Cao, Qingqian Tu, Sheng Guo, Bo Zheng, Huiyuan Liu, and Yang Yang. Policysim: An llm-based agent social simulation sandbox for proactive policy optimization. In *Proceedings of the ACM Web Conference 2026*, pp. 4781–4792, 2026.

Long Jin, Yang Chen, Tianyi Wang, Pan Hui, and Athanasios V Vasilakos. Understanding user behavior in online social networks: A survey. *IEEE communications magazine*, 51(9):144–150, 2013.

Haewoon Kwak, Changhyun Lee, Hosung Park, and Sue Moon. What is twitter, a social network or a news media? In *Proceedings of the 19th international conference on World wide web*, pp. 591–600, 2010.

Cheng Li, Jiaqi Ma, Xiaoxiao Guo, and Qiaozhu Mei. Deepcas: An end-to-end predictor of information cascades. In *Proceedings of the 26th international conference on World Wide Web*, pp. 577–586, 2017.

Peng Lu, Shizhao Nie, Zheng Wang, Ziwei Jing, Jianwu Yang, Zhongxiang Qi, and Wangmo Pujia. Predicting durations of online collective actions based on peaks' heights. *Communications in Nonlinear Science and Numerical Simulation*, 55:338–354, 2018a.

Peng Lu, Zheng Wang, Shizhao Nie, Wangmo Pujia, Pengfei Lu, and Baosheng Chen. Exploring the participate propensity in cyberspace collective actions: The 5‰ rule. *Physica A: Statistical Mechanics and Its Applications*, 503:582–590, 2018b.

Niklas Luhmann. *Social systems*. stanford university Press, 1995.

Hongyuan Mei and Jason M Eisner. The neural hawkes process: A neurally self-modulating multivariate point process. *Advances in neural information processing systems*, 30, 2017.

Qirui Mi, Mengyue Yang, Xiangning Yu, Zhiyu Zhao, Cheng Deng, Bo An, Haifeng Zhang, Xu Chen, and Jun Wang. Mf-llm: Simulating population decision dynamics via a mean-field large language model framework. *arXiv preprint arXiv:2504.21582*, 2025.

Shizhao Nie, Zheng Wang, Wangmo Pujia, Yuan Nie, and Peng Lu. Big data prediction of durations for online collective actions based on peak's timing. *Physica A: Statistical Mechanics and its Applications*, 492:138–154, 2018.

Daniele Notarmuzi, Claudio Castellano, Alessandro Flammini, Dario Mazzilli, and Filippo Radicchi. Universality, Criticality and Complexity of Information Propagation in Social Media. *Nature Communications*, 13(1):1308, 2022.

Joon Sung Park, Lindsay Popowski, Carrie Cai, Meredith Ringel Morris, Percy Liang, and Michael S Bernstein. Social simulacra: Creating populated prototypes for social computing systems. In *Proceedings of the 35th annual ACM symposium on user interface software and technology*, pp. 1–18, 2022.

Joon Sung Park, Joseph O'Brien, Carrie Jun Cai, Meredith Ringel Morris, Percy Liang, and Michael S Bernstein. Generative Agents: Interactive Simulacra of Human Behavior. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*, pp. 1–22, 2023.

Joon Sung Park, Carolyn Q. Zou, Aaron Shaw, Benjamin Mako Hill, Carrie Cai, Meredith Ringel Morris, Robb Willer, Percy Liang, and Michael S. Bernstein. Generative agent simulations of 1,000 people, 2024. URL https://arxiv.org/abs/2411.10109.

Jinghua Piao, Yuwei Yan, Jun Zhang, Nian Li, Junbo Yan, Xiaochong Lan, Zhihong Lu, Zhiheng Zheng, Jing Yi Wang, Di Zhou, et al. Agentsociety: Large-scale simulation of llm-driven generative agents advances understanding of human behaviors and society. *arXiv preprint arXiv:2502.08691*, 2025.

Zikai Song, Xiajie Li, Yunyao Zhang, Xinglang Zhang, Wei Yang, and Junqing Yu. Social intelligence modeling: A comprehensive survey from social perception to social simulation. ResearchGate preprint, 2026. URL https://doi.org/10.13140/RG.2.2.21157.87528. Preprint available on ResearchGate.

Flaminio Squazzoni, Wander Jager, and Bruce Edmonds. Social Simulation in the Social Sciences: A Brief Overview. *Social Science Computer Review*, 32(3):279–294, 2014.

Michael Szell, Sébastian Grauwin, and Carlo Ratti. Contraction of online response to major events. *PLoS one*, 9(2):e89052, 2014.

Ziyi Yang, Zaibin Zhang, Zirui Zheng, Yuxian Jiang, Ziyue Gan, Zhiyu Wang, Zijian Ling, Jinsong Chen, Martz Ma, Bowen Dong, Prateek Gupta, Shuyue Hu, Zhenfei Yin, Guohao Li, Xu Jia, Lijun Wang, Bernard Ghanem, Huchuan Lu, Chaochao Lu, Wanli Ouyang, Yu Qiao, Philip Torr, and Jing Shao. Oasis: Open agent social interaction simulations with one million agents, 2025. URL https://arxiv.org/abs/2411.11581.

Xinnong Zhang, Jiayu Lin, Xinyi Mou, Shiyue Yang, Xiawei Liu, Libo Sun, Hanjia Lyu, Yihang Yang, Weihong Qi, Yue Chen, et al. Socioverse: A world model for social simulation powered by llm agents and a pool of 10 million real-world users. *arXiv preprint arXiv:2504.10157*, 2025a.

Yunyao Zhang, Zikai Song, Hang Zhou, Wenfeng Ren, Yi-Ping Phoebe Chen, Junqing Yu, and Wei Yang. $ga - s^3$: Comprehensive social network simulation with group agents. In Wanxiang Che, Joyce Nabende, Ekaterina Shutova, and Mohammad Taher Pilehvar (eds.), *Findings of the Association for Computational Linguistics: ACL 2025*, pp. 8950–8970, Vienna, Austria, July 2025b. Association for Computational Linguistics. ISBN 979-8-89176-256-5. doi: 10.18653/v1/2025.findings-acl.468. URL https://aclanthology.org/2025.findings-acl.468/.

Yunyao Zhang, Yihao Ai, Zuocheng Ying, Qirui Mi, Junqing Yu, Wei Yang, and Zikai Song. Coupling macro dynamics and micro states for long-horizon social simulation. *arXiv preprint arXiv:2604.05516*, 2026a.

Yunyao Zhang, Zuocheng Ying, Xinglang Zhang, Junqing Yu, Peng Fang, Xu Chen, Wei Yang, and Zikai Song. Intervensim: Intervention-aware social network simulation for opinion dynamics. *arXiv preprint arXiv:2604.06600*, 2026b.

Qingyuan Zhao, Murat A Erdogdu, Hera Y He, Anand Rajaraman, and Jure Leskovec. Seismic: A self-exciting point process model for predicting tweet popularity. In *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 1513–1522, 2015.

Appendix

A. Related Work

B. Theoretical Foundations of the Simulation Target

C. Derivations of the Simulation Method

D. Detailed Experimental Setup

E. Additional Experimental Results and Analysis

A Related Work

We review four research strands in their broad historical progression, through a common boundary: *which variables are predicted, which are supplied as external conditions, and at what level the simulated state is defined*. For the task considered here, future participation and communication are prediction targets, external events are declared conditions, and opinion trajectories are observables derived from communication. The distinctions below concern mismatches with this task, rather than a blanket rejection of earlier simulation objectives.

Diffusion Forecasting: Predicting Activity without Generating Language

Hawkes processes model event arrivals through history-dependent intensities (Hawkes, 1971), with subsequent work distinguishing endogenous amplification from responses to external shocks (Crane & Sornette, 2008). *SEISMIC* predicts eventual popularity from early reshares, while *DeepCas* learns future cascade growth from observed diffusion structure (Zhao et al., 2015; Li et al., 2017). These methods explicitly predict future activity and evaluate it against held-out observations. Neural point processes extend this line to typed, timestamped event generation, and *HawkesTopic* incorporates textual topics into diffusion modeling (Mei & Eisner, 2017; He et al., 2015). The prediction target nevertheless **stops short of the communication process**: event times, cascade sizes, and topic representations do not specify what participants will say or how their expressed opinions will evolve. *This strand forecasts social activity while leaving its language-mediated development ungenerated.*

LLM Social Sandboxes: Persistent Agents without Participation Forecasting

LLMs introduce natural-language memory, planning, and interaction into agent-based simulation, exemplified by *Generative Agents* (Park et al., 2023). *OASIS* and *AgentSociety* scale this paradigm to larger social environments (Yang et al., 2025; Piao et al., 2025). *Their central modeling unit is a persistent agent: a population is instantiated, and its members repeatedly interact and update internal states*. Open online events pose a different structural problem. Participation is sparse and time-varying: accounts enter, disappear, and return, while future contributors need not belong to the initial population. **Who appears at each timestep is a prediction target**, not merely a scheduling choice within an existing cast. *Generating activity among persistent agents leaves event-specific participation forecasting unresolved, including its coupling to future communications and observable opinion trajectories.*

Mean-Field Simulators: Rendering Content on Supplied Participation

Mean-field LLM simulators pursue quantitative trajectory alignment by coupling individual decisions with evolving population signals. *MF-LLM* learns a textual mean field, while *M^{2}Sim* introduces latent micro states and long-horizon consistency training (Mi et al., 2025; Zhang et al., 2026a). Their response-generation loop operates on the active set supplied at each step; M^{2}Sim explicitly takes active-pool sizes and descriptors as inputs. Under realized-participation conditioning, this interface **places part of the future answer on the input side**: the simulator generates what the supplied participants say, while their presence and number are already given. *Agreement with real semantic distributions then measures content generation conditional on participation, not prediction of participation itself.* Stronger macro–micro coupling cannot remove this dependency. Without participation prediction, the system remains a content renderer along a supplied trajectory rather than a joint forecaster of activity and communication.

INTERVENTION-AWARE SIMULATION: CONDITIONING VS. ENDOGENIZING EXTERNAL CONDITIONS

Recent social simulators increasingly incorporate interventions into computational experiments. *Social Simulacra* treats interventions as experimenter-specified conditions (Park et al., 2022), whereas *PolicySim* and *IntervenSim* endogenize intervention selection within the simulated dynamics (Huang et al., 2026; Zhang et al., 2026b). Such formulations support intervention optimization, but differ from conditional social forecasting when the external-event stream is specified: **internally selecting or modifying an intervention changes the prediction target**, replacing the declared condition with a generated scenario. Moreover, aggregate or group-level responses do not generate the underlying account-level participation and communication. *Conditional social simulation therefore requires both adherence to specified external events and generation of the communications through which collective dynamics emerge*.

Across these strands, language is omitted from the forecast, participation is confined to an instantiated cast, future structure is supplied as input, or external conditions are generated inside an aggregated simulator. The missing formulation follows a single boundary: **starting from observed history, condition on external events, predict participation and communication, and evaluate their induced observable trajectories**.

B THEORETICAL FOUNDATIONS OF THE SIMULATION TARGET

This section formalizes two requirements underlying the communication-centered formulation: the simulation target must be grounded in observable records, and participation must remain part of the predicted process.

B.1 AN IDENTIFICATION LIMIT FOR UNOBSERVED INDIVIDUAL DYNAMICS

Let O collect the available platform records. Consider two latent-process models P_0 and P_1, with observable laws P_0^O and P_1^O. Let $\tau_j = \tau(P_j)$ denote a scalar property of the hidden individual dynamics, such as the probability of a state change during inactivity, and write $\Delta = |\tau_1 - \tau_0|$.

Proposition B.1 (Observation-limited identification)**.** *For any estimator $\widehat{\tau}(O)$ based only on observable records,*

$$\inf_{\widehat{\tau}} \max_{j\in\{0,1\}} \mathbb{E}_{P_j^O}\left[\left(\widehat{\tau}(O) - \tau_j\right)^2\right] \geq \frac{\Delta^2}{4}\left(1 - \mathrm{TV}(P_0^O, P_1^O)\right), \tag{12}$$

where TV *denotes total variation distance. In particular, observationally equivalent models with different latent dynamics imply an irreducible worst-case error of at least $\Delta^2/4$.*

Proof. Let p_0 and p_1 be densities of the observable laws with respect to a common dominating measure μ. For an arbitrary estimator, define

$$R_j = \mathbb{E}_{P_j^O}\left[\left(\widehat{\tau}(O) - \tau_j\right)^2\right].$$

Then

$$\begin{aligned} R_0 + R_1 &= \int \left[p_0(o)(\widehat{\tau}(o) - \tau_0)^2 + p_1(o)(\widehat{\tau}(o) - \tau_1)^2\right] d\mu(o) \\ &\geq \int \min\{p_0(o), p_1(o)\}\left[(\widehat{\tau}(o) - \tau_0)^2 + (\widehat{\tau}(o) - \tau_1)^2\right] d\mu(o) \\ &\geq \frac{\Delta^2}{2}\int \min\{p_0(o), p_1(o)\}\, d\mu(o) \\ &= \frac{\Delta^2}{2}\left(1 - \mathrm{TV}(P_0^O, P_1^O)\right). \end{aligned} \tag{13}$$

The second inequality follows from

$$(a - \tau_0)^2 + (a - \tau_1)^2 = 2\left(a - \frac{\tau_0 + \tau_1}{2}\right)^2 + \frac{\Delta^2}{2}.$$

Using $\max\{R_0, R_1\} \geq (R_0+R_1)/2$ and taking the infimum over estimators proves Eq. equation 12. □

Sparse participation admits precisely this ambiguity. Suppose an account is observed only at steps 0 and 2, with observations revealing its state at those steps. The hidden paths $(0,0,0)$ and $(0,1,0)$ produce identical observable records but imply different values of $\tau(P) = P(Z_1 \neq Z_0)$. Here $\Delta = 1$ and $P_0^O = P_1^O$, giving a lower bound of $1/4$.

Increasing simulator capacity cannot resolve a distinction absent from the observations. The bound concerns recovery of hidden individual dynamics, not prediction of subsequent communication. This motivates defining the empirical target through observable events, while using participant-level latent representations as conditional computational features rather than identified psychological trajectories.

B.2 The Prediction Burden Removed by Participation Conditioning

Fix a declared external stream $e_{1:T}$ and a common initial history. Let P^e and Q^e denote the reference and simulator laws of the observable rollout. At each step, write

$$X_t = (\mathcal{H}_{t-1}, e_t), \qquad S_t = \mathcal{U}_t,$$

and let C_t collect the participant-indexed actions and contents. Thus $\mathcal{Y}_t$ is represented by (S_t, C_t). Both processes use the same deterministic update of the complete observable history.

Define the content-only discrepancy on reference histories as

$$\mathcal{D}_{\text{comm}} = \sum_{t=1}^{T} \mathbb{E}_{P^e}\left[\mathrm{KL}\left(p(C_t \mid S_t, X_t) \,\|\, q(C_t \mid S_t, X_t)\right)\right]. \tag{14}$$

Proposition B.2 (Sequential participation–communication decomposition)**.** *Assume $P^e \ll Q^e$ and finite divergences. The full rollout discrepancy satisfies*

$$\mathrm{KL}(P^e \| Q^e) = \mathcal{D}_{\text{part}} + \mathcal{D}_{\text{comm}}, \tag{15}$$

$$\mathcal{D}_{\text{part}} = \sum_{t=1}^{T} \mathbb{E}_{P^e}\left[\mathrm{KL}\left(p(S_t \mid X_t) \,\|\, q(S_t \mid X_t)\right)\right]. \tag{16}$$

Consequently, $\mathrm{KL}(P^e \| Q^e) \geq \mathcal{D}_{\text{comm}}$, with equality exactly when the participation kernels agree on reference histories almost surely.

Proof. Applying the probability chain rule over time and within each step gives

$$\begin{aligned}\log \frac{P^e(\mathcal{Y}_{1:T})}{Q^e(\mathcal{Y}_{1:T})} &= \sum_{t=1}^{T} \log \frac{p(S_t, C_t \mid X_t)}{q(S_t, C_t \mid X_t)} \\ &= \sum_{t=1}^{T} \left[\log \frac{p(S_t \mid X_t)}{q(S_t \mid X_t)} + \log \frac{p(C_t \mid S_t, X_t)}{q(C_t \mid S_t, X_t)}\right].\end{aligned} \tag{17}$$

Taking expectation under P^e and conditioning on X_t for the first term and (S_t, X_t) for the second yields

$$\begin{aligned}\mathrm{KL}(P^e \| Q^e) = &\sum_{t=1}^{T} \mathbb{E}_{P^e}\left[\mathrm{KL}\left(p(S_t \mid X_t) \,\|\, q(S_t \mid X_t)\right)\right] \\ &+ \sum_{t=1}^{T} \mathbb{E}_{P^e}\left[\mathrm{KL}\left(p(C_t \mid S_t, X_t) \,\|\, q(C_t \mid S_t, X_t)\right)\right].\end{aligned} \tag{18}$$

Non-negativity of conditional KL gives the inequality. Equality holds if and only if every participation divergence vanishes almost surely. □

Conditioning content evaluation on realized participants leaves an entire component of process prediction untested. Even when $\mathcal{D}_{\text{comm}} = 0$, the full rollout discrepancy can remain positive through $\mathcal{D}_{\text{part}}$. The decomposition concerns the joint event-process law, rather than the KL between aggregate attitude distributions used in empirical evaluation.

Algorithm 1 Open-Population Communication Simulation

Require: Global user pool $\mathcal{U}$, horizon T; volume forecaster f_ω, activation model q_ψ, state transition model f_η, communication policy π_ϕ, trajectory critic C_γ, textual updater g_ξ; environmental inputs $\{\mathcal{E}_t\}_{t=1}^T$

Initialize: platform history $\mathcal{H}_0$ and collective state (m_0, r_0)
for $t = 1$ **to** T **do**
  $\hat{N}_t \leftarrow f_\omega(\mathcal{H}_{t-1}, \mathcal{E}_t)$ ▷ Participation modeling
  $\mathcal{U}_t \leftarrow \textsc{Activate}(\mathcal{U}, \hat{N}_t; q_\psi, \mathcal{H}_{t-1}, \mathcal{E}_t)$
  $(\hat{m}_t, \{h_{t,j}\}) \leftarrow f_\eta(\mathcal{H}_{t-1}, \mathcal{E}_t, P_t)$ ▷ Mean-field transition
  for all $u_{t,j} \in \mathcal{U}_t$ **do**
    $y_{t,j} \leftarrow \textsc{GenerateSelect}(\pi_\phi, C_\gamma; \mathcal{H}_{t-1}, \mathcal{E}_t, p_{u_{t,j}}, h_{t,j}, \hat{m}_t)$ ▷ Future-constrained policy
  end for
  $\mathcal{Y}_t \leftarrow \{y_{t,j}\}$
  $m_t \leftarrow \textsc{MeanFieldUpdate}(m_{t-1}, \mathcal{Y}_t)$ ▷ Realized mean field
  $r_t \leftarrow g_\xi(r_{t-1}, m_t, \mathcal{Y}_t, \mathcal{E}_t)$
  $\mathcal{H}_t \leftarrow \textsc{Update}(\mathcal{H}_{t-1}, \mathcal{E}_t, \mathcal{Y}_t, m_t, r_t)$ ▷ History update
end for
Output: $\{\mathcal{Y}_t\}_{t=1}^T$

C Derivations of the Simulation Method

C.1 Participation and the Observable Transition Kernel

Write $X_t = (\mathcal{H}_{t-1}, \mathcal{E}_t)$ and $M = |\mathcal{U}|$. For a feasible predicted count $n = \hat{N}_t \in \{0, \dots, M\}$, define positive activation weights

$$w_i = q_\psi(u_i \mid p_i, X_t).$$

Sequential sampling without replacement induces the ordered activation law

$$Q_\psi^{\mathrm{ord}}(u_{i_1}, \dots, u_{i_n} \mid n, X_t) = \prod_{j=1}^{n} \frac{w_{i_j}}{\sum_{\ell \notin \{i_1, \dots, i_{j-1}\}} w_\ell}. \tag{19}$$

The probability of an unordered participant set S is

$$Q_\psi(S \mid n, X_t) = \sum_{\sigma \in \mathrm{Perm}(S)} Q_\psi^{\mathrm{ord}}(\sigma \mid n, X_t). \tag{20}$$

Given S, the state model produces $(\hat{m}_t, \{h_{t,j}\}) = f_\eta(X_t, P_t)$. Let

$$x_{t,j} = (X_t, p_{u_{t,j}}, h_{t,j}, \hat{m}_t)$$

be the communication context, and let $\bar{\pi}_{\phi,\gamma}$ denote the distribution induced by candidate generation and selection. With independent candidate sets across activated participants, the complete step kernel is

$$Q_\theta(n, S, C_t \mid X_t) = Q_\omega(n \mid X_t)\, Q_\psi(S \mid n, X_t) \prod_{j=1}^{n} \bar{\pi}_{\phi,\gamma}(a_{t,j}, c_{t,j} \mid x_{t,j}), \tag{21}$$

where the point forecast corresponds to $Q_\omega(n \mid X_t) = \mathbf{1}\{n = f_\omega(X_t)\}$.

Proposition C.1 (Normalized open-population transition)**.** *If Q_ω is supported on $\{0, \dots, M\}$ and each selected communication law is normalized, Eq. equation 21 defines a normalized transition with exactly n distinct active accounts.*

Proof. At each activation draw, the conditional probabilities sum to one over the remaining accounts. Successively summing Eq. equation 19 over its final, penultimate, and preceding draws therefore gives one. Grouping permutations by their underlying sets preserves normalization. Consequently,

$$\sum_{n=0}^{M} \sum_{S:|S|=n} \sum_{C_t} Q_\theta(n, S, C_t \mid X_t) = \sum_{n=0}^{M} Q_\omega(n \mid X_t) \sum_{S:|S|=n} Q_\psi(S \mid n, X_t)$$
$$= 1. \tag{22}$$

For $n = 0$, the participant set and communication collection are empty and the empty product equals one. Exclusion operates within each step, allowing an account to reappear at later steps. □

For $n > 0$, let $A(y) \in \{e_1, e_2, e_3\}$ encode the expressed attitude of communication y. The realized state update is

$$m_t = \frac{1}{n}\sum_{j=1}^{n} A(y_{t,j}), \qquad r_t = g_\xi(r_{t-1}, m_t, \mathcal{Y}_t, \mathcal{E}_t). \tag{23}$$

Thus $\hat{m}_t$ **conditions generation, while** m_t **records its realized outcome** and enters the next-step history.

C.2 Differentiable Future-Constrained Learning

For a training communication $y^\star$ and context x, let $v_\phi(x, y^\star)$ collect the policy's teacher-forced prompt and response features. The trajectory critic acts on these continuous features:

$$\widehat{\mathbf{m}}(v) = C_\gamma(x, v), \qquad d_\gamma(v) = \sum_{k:b_{t,k}=1} D\big(m^\star_{t+k}, [\widehat{\mathbf{m}}(v)]_k\big), \tag{24}$$

where $b_{t,k}$ indicates an observed future target. An empty valid-horizon set gives $d_\gamma(v) = 0$.

Let v_+ be the matched feature vector and $\{v_r^-\}_{r=1}^R$ its masked or shuffled counterparts. Define

$$\begin{aligned} B_t &= \mathbf{1}\left\{\sum_k b_{t,k} > 0\right\}, \\ a_r &= \delta + d_\gamma(v_+) - d_\gamma(v_r^-), \\ \mathcal{L}_{\text{sens}} &= \frac{B_t}{R}\sum_{r=1}^{R}[a_r]_+. \end{aligned} \tag{25}$$

For a valid sample, identical matched and negative predictions incur loss δ. Zero sensitivity loss requires

$$d_\gamma(v_r^-) \geq d_\gamma(v_+) + \delta \qquad \text{for every } r.$$

The joint objective in Eq. equation 11 becomes

$$\mathcal{L}_{\text{joint}} = \mathcal{L}_{\text{comm}} + \lambda_{\text{traj}} d_\gamma(v_+) + \lambda_{\text{sens}} \frac{B_t}{R}\sum_{r=1}^{R}[a_r]_+. \tag{26}$$

Gradient transmission to the policy. Write $J_+ = \partial v_+/\partial\phi$, $J_r^- = \partial v_r^-/\partial\phi$, $g_+ = \nabla_{v_+} d_\gamma(v_+)$, and $g_r^- = \nabla_{v_r^-} d_\gamma(v_r^-)$. Away from hinge points, the chain rule yields

$$\begin{aligned} \nabla_\phi \mathcal{L}_{\text{joint}} = \nabla_\phi \mathcal{L}_{\text{comm}} &+ \lambda_{\text{traj}} J_+^\top g_+ \\ &+ \frac{\lambda_{\text{sens}} B_t}{R}\sum_{r=1}^{R}\mathbf{1}\{a_r > 0\}\left(J_+^\top g_+ - (J_r^-)^\top g_r^-\right). \end{aligned} \tag{27}$$

The critic gradient is

$$\begin{aligned} \nabla_\gamma \mathcal{L}_{\text{joint}} = \lambda_{\text{traj}} \nabla_\gamma d_\gamma(v_+) &\\ &+ \frac{\lambda_{\text{sens}} B_t}{R}\sum_{r=1}^{R}\mathbf{1}\{a_r > 0\}\left(\nabla_\gamma d_\gamma(v_+) - \nabla_\gamma d_\gamma(v_r^-)\right). \end{aligned} \tag{28}$$

At $a_r = 0$, a hinge subgradient is used. Constant masked components have zero Jacobian; shuffled features retain their reindexed gradient paths.

Equations equation 27–equation 28 establish the coupling: **future-trajectory supervision updates both the critic and the policy through continuous communication features**. Observed tokens remain fixed during teacher forcing, while their hidden representations depend on the policy parameters.

Table 4: **Information coverage of social-simulation benchmarks.** Traffic: event-time aggregate activity; Micro: individual communications; Reversal: opinion-reversal events. The final three columns indicate structured intervention annotations at each level. ✓: explicitly covered; ×: not specified as a benchmark component.

	Observed social dynamics			Intervention annotations		
Benchmark	**Traffic**	**Micro**	**Reversal**	**Event**	**Platform**	**Participant**
GA-S^3 (Zhang et al., 2025b)	✓	×	×	×	×	×
MF-LLM (Mi et al., 2025)	×	✓	×	×	×	×
M^{2}Sim (Zhang et al., 2026a)	×	✓	✓	×	×	×
This work	✓	✓	✓	✓	✓	✓

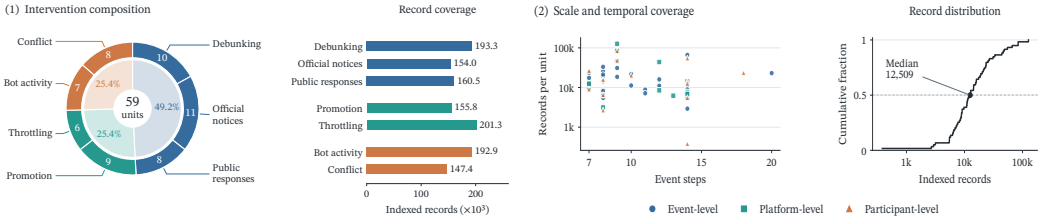

Figure 4: **Dataset composition and scale. (1)** Inner and outer rings show unit counts across three intervention levels and seven types, respectively. Matching colors identify types in the adjacent bars, which report indexed records in thousands. Legend percentages denote unit shares by intervention level. **(2)** Each point shows one unit's duration and indexed-record count; the empirical cumulative distribution summarizes record counts across units (median: 12,509). Record-count axes in (2) are logarithmic.

C.3 Future-Blind Lookahead and Candidate Selection

Let $\mathcal{I}_t$ contain the frozen models, admissible participant information, simulated history, and environmental inputs declared available at step t. The state-model lookahead produces a reference trajectory

$$\widetilde{\mathbf{m}}_{t+1:t+K} = R_\eta(\mathcal{I}_t, S_t, \hat{m}_t; \zeta_t), \tag{29}$$

where ζ_t is an inference seed. All quantities used in the lookahead are declared inputs or internally generated predictions.

For a nonempty candidate set $\mathcal{C}_{t,j} = \{y^{(1)}, \ldots, y^{(J)}\}$, selection uses

$$\begin{aligned} s_t(y) &= D_K(C_\gamma(x_{t,j}, v_\phi(x_{t,j}, y)), \widetilde{\mathbf{m}}_{t+1:t+K}), \\ \widehat{y}_{t,j} &\in \arg\min_{y \in \mathcal{C}_{t,j}} s_t(y), \end{aligned} \tag{30}$$

with a fixed tie-breaking rule. Here D_K aggregates trajectory discrepancies across predicted horizons.

Proposition C.2 (Future-blind rollout)**.** *If every inference module depends only on $\mathcal{I}_t$ and inference randomness, two runs with identical admissible inputs, frozen parameters, and random seeds produce identical rollouts regardless of the withheld future records.*

Proof. At the first step, both runs produce identical counts, participant sets, state predictions, candidate sets, and lookahead references. Equation equation 30 therefore selects identical communications, yielding identical updated histories. Induction gives equality at every subsequent step. □

Proposition C.3 (Finite-candidate selection bound)**.** *Let $s_t^\star$ be a target compatibility score on $\mathcal{C}_{t,j}$. If $\sup_{y \in \mathcal{C}_{t,j}} |s_t(y) - s_t^\star(y)| \le \epsilon$, then*

$$s_t^\star(\widehat{y}_{t,j}) - \min_{y \in \mathcal{C}_{t,j}} s_t^\star(y) \le 2\epsilon. \tag{31}$$

Table 5: **Simulation settings and participant pools.** Warm-up is measured in interactions and content length in tokens. Replacement refers to sampling within a simulation step.

Generation and candidate selection

Warm-up	Seed	Candidates J	Temperature	Top-p
50	46	4	0.8	0.9
Repetition penalty	**Content limit**	**MC-dropout samples**	**Uncertainty weight**	**Quality weight**
1.05	48	4	0.10	0.05

Participant-pool comparison

Participant pool	Size	Sampling unit	Sampling	Replacement
Event-specific	Event-dependent	User account	Uniform	No
External (SocioVerse)	15,992	User account	Uniform	No

Proof. For $y^\star \in \arg\min_{y \in \mathcal{C}_{t,j}} s_t^\star(y)$,

$$s_t^\star(\widehat{y}_{t,j}) \le s_t(\widehat{y}_{t,j}) + \epsilon \le s_t(y^\star) + \epsilon \le s_t^\star(y^\star) + 2\epsilon.$$

□

The first result establishes the inference information boundary. The second quantifies selection accuracy within the generated candidate set under bounded scoring error. Together, they separate **future-supervised learning** from **future-blind generation and selection**.

D Detailed Experimental Setup

D.1 Dataset and Preprocessing

The dataset combines **traffic dynamics, individual communications, opinion reversals, and timed interventions** across **59 intervention-event units**, comprising 1,205,150 indexed interaction records and 606 discrete event steps. Each unit links communication texts and participant features to interaction counts, attitude trajectories, and intervention timing; different intervention contexts of the same underlying topic form separate units. Table 4 compares benchmark coverage, and Figure 4 summarizes the dataset's composition and scale.

Intervention taxonomy. Seven intervention types are organized by **where they enter the social process**. **Event-level** interventions update event information: *debunking* corrects circulating claims, *official notices* provide institutional statements, and *public responses* capture explanations from involved parties. **Platform-level** interventions regulate exposure through *promotion* or *throttling*. **Participant-level** interventions comprise *bot activity*, introducing automated participation, and *conflict*, capturing antagonistic interactions. Their occurrence steps are aligned with the communication and attitude trajectories.

Preprocessing and supervision. Training, validation, and test contain **44, 8, and 7 intervention-event units**, respectively. Training and validation yield 27,757 and 5,270 policy examples; the test set spans five underlying topics, 283,952 indexed records, and 68 event steps. Each example conditions on available event context, participant features, the preceding mean-field synopsis, and an activation-local representation. The target consists of a `comment` or `share` action and its text. Attitude distributions use the order $[\text{positive}, \text{neutral}, \text{negative}]$. Future supervision comprises cumulative attitude distributions over the next $K = 10$ steps, with unavailable horizons masked.

D.2 Simulation Protocol

Initialization and rollout. Each event retains its post context and begins with a 50-interaction warm-up, followed by participant sampling and policy generation. Selected communications update the simulated history and mean field. Table 5 summarizes the settings.

Table 6: **Key model and training hyperparameters.** Hyperparameters are shared across backbone configurations unless otherwise specified. Epoch counts describe the initial training schedule before validation-based checkpoint selection.

State model		Communication policy		Trajectory critic	
Hidden dimension	192	LoRA rank / alpha	32 / 32	Hidden dimension	256
Layers / heads	2 / 6	Sequence length	704	Recurrent layers	1
Latent dimension	64	Effective batch size	32	Dropout	0.2
Participants / step	128	SFT learning rate	8×10^{-6}	Future horizon K	10
Learning rate	10^{-3}	Joint learning rate	3×10^{-6}	MC-dropout samples	4
Epoch budget	12	SFT epochs	3	Warm-up learning rate	5×10^{-5}
Auxiliary horizon	3	Initial joint epochs	2	Joint learning rate	3×10^{-5}
Loss coefficients and weighting schedule					
Increment weight	0.35	Direction weight	0.15	Curvature weight	0.05
MC-consistency weight	0.01	Relative sensitivity weight	0.5	Sensitivity margin δ	0.05
Maximum future weight	0.15	Horizon discount	0.95	Future-weight ramp	20%

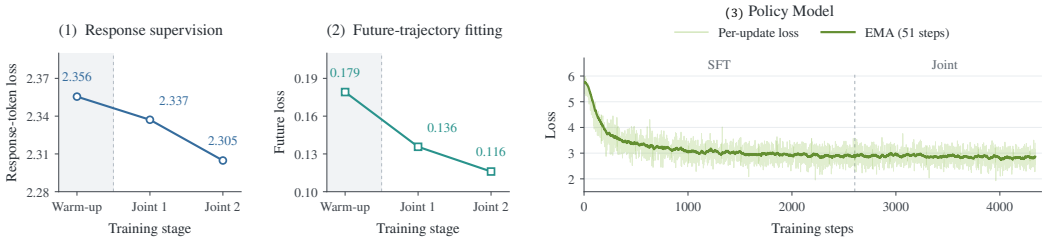

Figure 5: **Training dynamics of Qwen2.5-1.5B. (1) Response:** epoch-averaged response-token loss. **(2) Future:** epoch-averaged future-fitting objective $\mathcal{L}_{\text{fit}}$. Both panels cover critic warm-up and two joint-training epochs. **(3) Policy:** stage-specific training loss over 2,604 SFT updates and 1,736 joint-training updates. The light curve shows the mean mini-batch loss per update; the dark curve shows an exponential moving average with a span of 51 updates. The dashed line marks the SFT–joint boundary, where smoothing restarts.

Participant pools. OpenSim uses an *event-specific participant pool* constructed from the corresponding event's observed participants, whereas OpenSim† uses a frozen *external participant pool* of 15,992 real-user profiles from SocioVerse (Zhang et al., 2025a). Both settings uniformly sample user accounts without replacement within each step, while allowing users to participate again in subsequent steps. Event context and stepwise communication counts are held fixed in the pool comparison. Neither setting consumes the observed future participant sequence.

Communication generation. At each policy call, four candidate communications are generated. Selection minimizes masked KL divergence to the state model's predicted future attitude trajectory, together with MC-dropout uncertainty and output-quality penalties.

D.3 Training Curves and Hyperparameters

Training configuration. Communication policies are adapted from Qwen2-1.5B, Qwen2.5-1.5B/3B/7B, and Qwen3-1.7B using LoRA. All backbones follow a three-stage framework: communication SFT, critic-only warm-up, and joint optimization. Warm-up freezes the policy and trains the trajectory critic; joint training updates both the policy adapters and critic. Checkpoint selection uses validation communication quality and response-sensitive trajectory prediction. Additional critic calibration, when required by validation, keeps the policy frozen.

Optimization. Table 6 reports the key model hyperparameters, training schedules, and loss coefficients. All modules are optimized with AdamW using random seed 42. The future-objective weight

Table 7: **Component ablations across backbones.** W/O ST OBJECTIVE removes state-transition supervision, while W/O FUTURE CONSTRAINT removes the future-constrained policy objective. **Bold**/underline indicate the best/second-best result within each backbone.

Model	Setting	KL ↓	Wass. ↓	DTW ↓	Macro-F1 ↑	Micro-F1 ↑	NLL ↓
Qwen 2-1.5B	**OpenSim**	**1.1352**	**0.1922**	**0.1261**	**0.4444**	**0.6616**	**4.3377**
	W/O ST OBJECTIVE	3.1716	0.3074	0.2644	0.4087	0.5079	4.4791
	W/O FUTURE CONSTRAINT	1.6934	0.2410	0.1914	0.4331	0.5597	4.4315
Qwen 2.5-1.5B	**OpenSim**	**1.0667**	**0.1724**	**0.1252**	**0.6097**	**0.6252**	4.5980
	W/O ST OBJECTIVE	2.9886	0.2725	0.2494	0.4222	0.5299	4.4942
	W/O FUTURE CONSTRAINT	1.5839	0.1940	0.1864	0.5369	0.5646	**4.3211**
Qwen 3-1.7B	**OpenSim**	**1.0929**	**0.1780**	**0.1098**	**0.5445**	**0.7610**	4.4414
	W/O ST OBJECTIVE	2.8526	0.2775	0.2412	0.4782	0.5776	4.5111
	W/O FUTURE CONSTRAINT	1.7857	0.2320	0.1919	0.4413	0.6587	**4.0974**

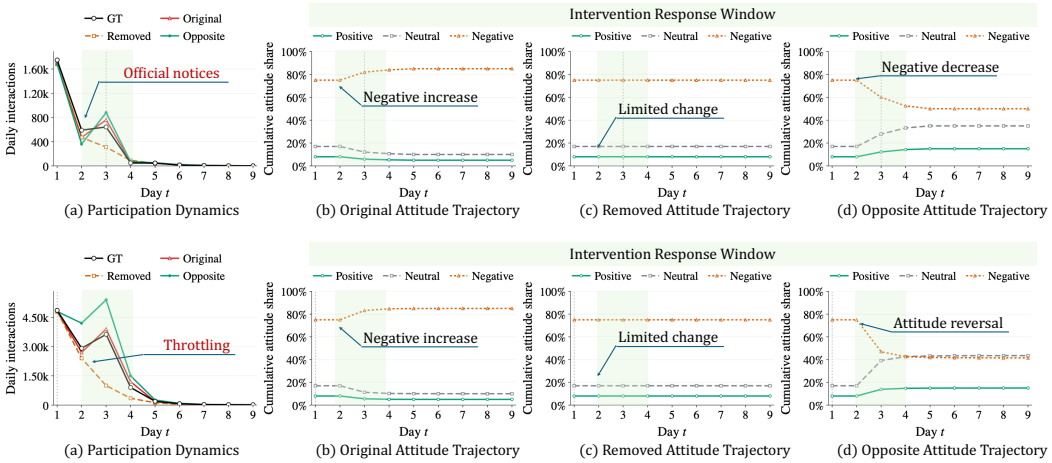

Figure 6: **Intervention responses in the Xi'an metro case.** The top and bottom rows examine official notices and throttling, respectively. Each row shows participation dynamics and cumulative attitude shares under ORIGINAL, REMOVED, and OPPOSITE. GT denotes observed participation; shaded regions mark intervention response windows.

increases linearly during the initial portion of joint training, while unavailable future horizons are masked throughout trajectory supervision.

Training dynamics. Figure 5 separates response supervision and future fitting from the overall policy optimization trace, showing component losses at the epoch level and policy loss at the optimizer-update level.

E ADDITIONAL EXPERIMENTAL RESULTS AND ANALYSIS

E.1 CASE STUDIES OF INTERVENTION RESPONSES

Figures 6 and 7 compare participation and cumulative attitude trajectories under ORIGINAL, REMOVED, and OPPOSITE. The original conditions assess reproduction of observed attitude trends, removal examines evolution without the corresponding intervention, and opposite conditions explore alternative governance responses.

In the Xi'an metro case, the initial response emphasized the passenger's disruption of public order, while public criticism focused on the security guard's conduct. To address this mismatch, the opposite official notice explicitly acknowledges improper handling, apologizes to the passenger, and announces investigation, accountability, and corrective measures. For throttling, the opposite condition removes additional exposure restrictions and allows discussion to spread naturally. These changes test whether addressing public concerns and maintaining open discussion can alleviate neg-

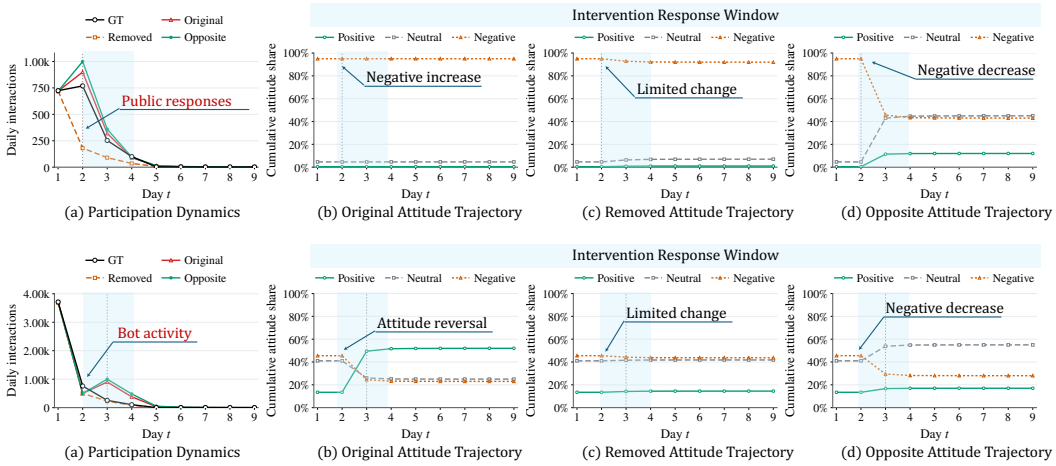

Figure 7: **Intervention responses in the Guilin tourism case.** The top and bottom rows examine public responses and bot activity, respectively. Each row shows participation dynamics and cumulative attitude shares under ORIGINAL, REMOVED, and OPPOSITE. GT denotes observed participation; shaded regions mark intervention response windows.

ative attitudes. The original conditions reproduce the observed persistence and further increase of negative attitudes. Removing the interventions weakens the participation response and leaves attitude shares comparatively stable. Under the opposite conditions, the negative share declines and the neutral share increases, with neutrality becoming slightly dominant when throttling is lifted. Increased participation therefore need not coincide with more negative attitudes.

In the Guilin tourism case, a tourist's complaint about dirty concert seats was followed by blocking by the official tourism account, extending the controversy from service quality to complaint handling. The subsequent official response did not resolve public dissatisfaction. The opposite response therefore emphasizes concrete remedies: unblocking and contacting the complainant, reporting cleaning inspections and complaint-resolution progress, and providing a follow-up feedback channel. This tests whether verifiable corrective measures can alleviate persistent dissatisfaction. The original response condition reproduces the observed dominance of negative attitudes. Removing the response produces little change, whereas the opposite response substantially reduces the negative share, with neutrality becoming the largest category. For bot activity, we replace positive stance promotion with neutral factual communication, restricting messages to verified information and updates on corrective actions. This comparison examines how the two communication strategies shape attitude evolution. The original condition reproduces the observed shift from negative to positive dominance, while the opposite condition yields a neutral-dominated trajectory. Together, these cases demonstrate reproduction of observed attitude trends and simulation of possible outcomes under alternative governance responses.

E.2 FINE-GRAINED SEMANTIC DISTRIBUTIONS

Figure 8 compares complete communication distributions across eight semantic dimensions. **The clearest differences concern how agents communicate, what they express, and how strongly their outputs concentrate on dominant categories.**

Behavior and communicative intent. Direct LLM and Social Retrieval allocate approximately 92% of their outputs to sharing, compared with 46.8% in the reference distribution. Their intent distributions are similarly concentrated on promotion, at 91.1% and 76.6%, respectively, against a reference share of 46.4%. OpenSim produces substantially more comments and opinions, bringing both distributions closer to the reference. The discrepancy therefore extends beyond action frequency: it also changes the semantic composition of the simulated communications.

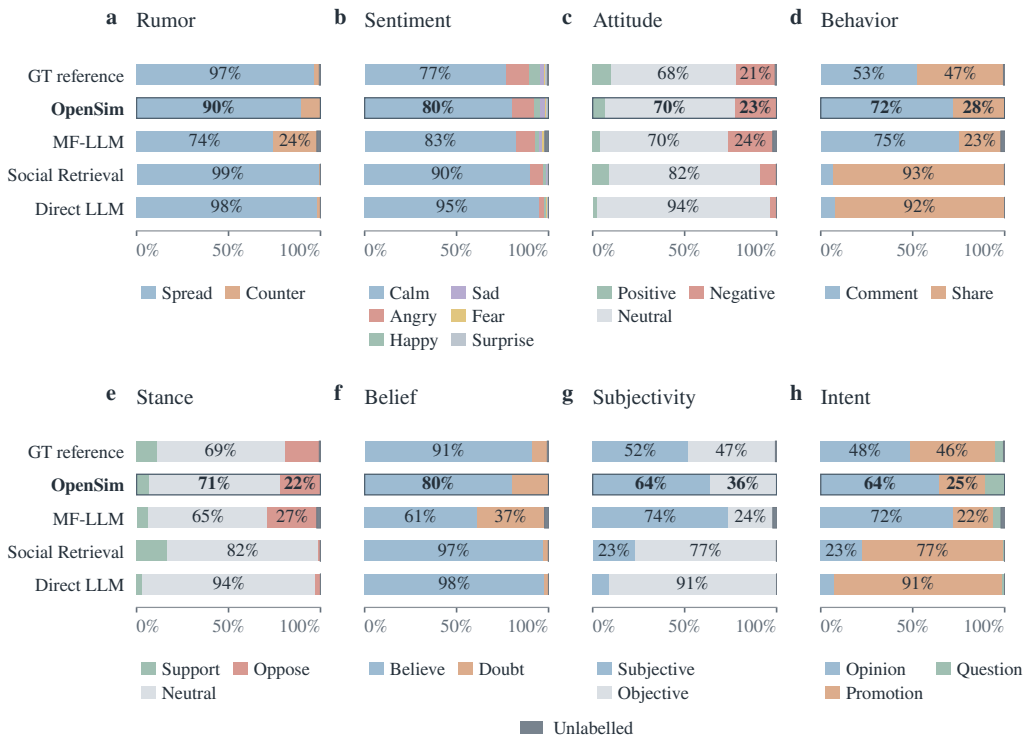

Figure 8: **Full semantic distributions across eight dimensions.** Qwen2-1.5B results pooled over seven test units and 68 event-day groups, with 10,879 record pairs per method. Each bar shows record-weighted category proportions; gray denotes unlabelled records. GT reference averages the real-label distributions across the four evaluation caches. Percentage labels are shown for segments of at least 20%.

Expression style and affective composition. OpenSim retains 63.7% subjective expression, compared with 52.0% in the reference, whereas Direct LLM and Social Retrieval produce predominantly objective outputs. A related concentration appears in emotion and attitude. The calm category accounts for 95.2% of Direct LLM outputs and 90.0% of Social Retrieval outputs, but 77.1% of the reference; OpenSim yields 80.3%. Its neutral-attitude share of 70.4% likewise remains close to the reference value of 68.2%. The stance distribution exhibits the same pattern, with OpenSim retaining substantial opposing responses alongside neutral communication. **These comparisons reveal differences that matching the dominant category alone would obscure.**

Rumor and belief. OpenSim assigns more probability mass to countering and doubt than Direct LLM and Social Retrieval. These categories also exceed their reference proportions, showing that broader response coverage does not automatically imply closer distributional alignment. The complete profiles distinguish this redistribution from the stronger reference agreement observed in behavior, expression style, and affective composition.