DECOUPLING OPENNESS AND CONNECTIVITY: NON-MONOTONIC EFFECTS IN LLM-BASED CULTURAL DYNAMICS

Anonymous authors
Paper under double-blind review

ABSTRACT

Cultural dynamics in multi-agent systems exhibit a counterintuitive phenomenon: local similarity-based interactions can lead to global fragmentation rather than convergence. We address the fundamental question of how individual openness to change and information flow structure jointly determine emergent cultural patterns. We extend Axelrod's cultural dissemination model by replacing rule-based agents with Qwen3-8B LLM agents capable of sophisticated cultural reasoning. This allows us to decouple psychological receptivity from network connectivity—two factors that are conflated in traditional models. Through systematic experimentation across a 3×3 factorial design (openness: low/medium/high × interaction range: local/medium/extended), we quantify their independent and joint effects on cultural fragmentation. Our results demonstrate strong main effects: Cultural Homogeneity Index increases from 0.279 to 0.437 with higher openness (1st order interactions, +57%), while optimal information flow (3rd order) achieves the highest convergence at 0.489 for high openness agents—representing 75% improvement over low openness baseline (0.279). Critically, we uncover a non-monotonic relationship where 3rd-order interactions consistently outperform both 1st and 5th-order across all openness levels, revealing an optimal balance between exploration and exploitation. Code can be found at https://anonymous.4open.science/r/YuLan-OneSim/.

1 INTRODUCTION

Figure 1: **Cultural Dynamics in Multi-Agent Systems: Main Results Overview.** This figure presents a comprehensive overview of our findings on how individual openness and information flow structure jointly influence cultural dynamics in multi-agent systems. The visualization demonstrates the key relationships between psychological factors (agent openness) and structural factors (information flow range) in determining cultural convergence versus fragmentation outcomes.

How do local cultural interactions produce global social patterns? Axelrod's seminal model Axelrod (1997) revealed a paradox: similarity-based interactions can drive polarization rather than convergence, fragmenting societies into distinct cultural regions. This counterintuitive phenomenon appears across domains—political echo chambers, residential segregation, organizational silos—motivating both theoretical understanding and practical interventions.

Traditional agent-based models face two critical limitations. First, they conflate psychological and structural factors: adoption propensity and interaction range are either fixed or coupled, preventing independent manipulation. This obscures a fundamental question: do openness and connectivity simply add, or do they interact in complex ways? Second, rule-based agents lack cognitive sophistication to capture realistic cultural reasoning. Simple stochastic rules ($P_{\text{adopt}} = f(\text{similarity})$) cannot represent how humans evaluate social influence—weighing contextual factors, considering identity trade-offs, and adapting strategies based on experience.

Recent advances in large language models enable a new approach. LLM agents can engage in contextual reasoning about cultural traits, evaluate social influence, and make adaptive decisions Park et al. (2023); Wang et al. (2024). This creates an opportunity to build more realistic simulations while maintaining experimental control, specifically enabling independent manipulation of psychological receptivity and structural connectivity.

This research makes the following key contributions:

• *Framework for decoupling factors*: We introduce a parametric extension decoupling individual openness (psychological receptivity) from information flow (interaction range), enabling systematic exploration through a 3×3 factorial design with 27 simulation runs.

• *Discovery of nonlinear dynamics*: Using Qwen3-8B agents, we reveal that medium openness exhibits accelerating convergence—a signature of cascade amplification—while information flow demonstrates a robust 3rd-order optimum contradicting monotonic connectivity assumptions.

• *Capacity-connectivity matching principle*: We discover non-additive interactions where optimal network structure depends fundamentally on psychological capacity. High openness achieves peak performance at 3rd-order (0.489), while both 1st-order (0.437) and 5th-order (0.433) underperform due to exposure bottlenecks and noise saturation, respectively.

Our findings establish quantitative relationships between micro-level parameters and macro-level patterns in cognitively realistic agents. Theoretically, we demonstrate that cultural dynamics emerge from complex psychological-structural interplay with nonlinear thresholds invisible to traditional models. Practically, we inform the design of social platforms and integration policies by identifying when local versus global connectivity promotes cohesion.

2 Related Work

2.1 Multi-Agent Interaction Dynamics

Classical models couple similarity-based interaction with state alignment Barbosa & Fontanari (2009). Extensions modify interaction rules through agreement thresholds MacCarron et al. (2020) and antagonistic features Gracia-Lázaro et al. (2021). However, these approaches directly tie interaction probability to similarity, lacking independent control over agent receptivity to dissimilar states.

2.2 Network Topology and Information Flow

Information propagation has been controlled through network structure and external signals. Broadcasting mechanisms can destabilize equilibria or induce global convergence Peres & Fontanari (2009); Rodríguez & Moreno (2010). Dynamic rewiring couples topology evolution with state updates Gracia-Lazaro et al. (2009), while fully-connected graphs provide analytical tractability Pinto & Balenzuela (2020). These methods typically fix local interaction rules while varying connectivity patterns, or introduce exogenous information sources rather than controllable spatial interaction ranges.

Phase transitions have been extensively mapped as functions of system parameters including state dimensionality, trait cardinality, and network topology Stivala et al. (2014); Barbosa & Fontanari (2009), with mean-field approximations yielding tractable phase diagrams Pedraza et al. (2020). However, existing characterizations do not systematically explore the joint parameter space of agent receptivity and spatial interaction scale.

2.3 LLM-BASED SOCIAL SIMULATION

Recent advances have enabled AI agents with sophisticated reasoning capabilities that simulate human-like behavior Xu et al. (2024). Unlike traditional rule-based agents, LLM-based agents engage in complex reasoning, adapt behavior based on context, and exhibit emergent learning patterns mirroring human cognitive processes. Our approach leverages Qwen3-8B Yang et al. (2025) to create agents capable of nuanced cultural reasoning, enabling systematic exploration of psychological-structural interactions that remain hidden when these parameters are structurally coupled.

3 METHOD

We present a novel framework for cultural dissemination that replaces traditional rule-based agents with large language model agents capable of sophisticated reasoning. Our approach introduces parametric control over two previously conflated factors—individual openness and information flow—enabling systematic investigation of their joint effects on emergent cultural patterns.

3.1 PROBLEM FORMULATION

Cultural state space. Consider a population of N agents arranged on a two-dimensional lattice $\mathcal{L} = \{(x, y) : 1 \leq x, y \leq \sqrt{N}\}$. Each agent i at position (x_i, y_i) possesses a cultural state vector $\mathbf{T}_i = (t_{i1}, \ldots, t_{iD})$ where $D = 5$ represents the number of cultural dimensions (music preference, culinary taste, fashion style, political orientation, leisure activity). Each trait takes values in a discrete space $\mathcal{V} = \{0, 1, \ldots, q-1\}$ with $q = 10$ possible values per dimension.

Cultural similarity. The similarity between agents i and j is the proportion of shared traits:

$$s_{ij} = \frac{1}{D} \sum_{d=1}^{D} \mathbb{1}[t_{id} = t_{jd}] \in [0, 1] \tag{1}$$

where $\mathbb{1}[\cdot]$ is the indicator function. This metric captures feature overlap: $s_{ij} = 0$ indicates complete cultural difference, $s_{ij} = 1$ indicates identity, and intermediate values represent partial overlap.

Interaction neighborhoods. We define interaction neighborhoods through spatial proximity parameterized by order k:

$$N_k(i) = \{j \in \mathcal{L} : d(i, j) \leq k\} \tag{2}$$

where $d(i, j) = |x_i - x_j| + |y_i - y_j|$ is the Manhattan (L1) distance on the lattice. For $k = 1$, $N_1(i)$ contains the four orthogonally adjacent neighbors (north, south, east, west). For $k = 3$, $N_3(i)$ includes all agents within 3 steps, encompassing diagonals and 2-hop connections ($|N_3(i)| \approx 28$ for interior agents). For $k = 5$, $N_5(i)$ spans a broad spatial range ($|N_5(i)| \approx 60$ for interior agents).

Dynamics. At each timestep t:

1. Each agent i selects one interaction partner $j \in N_k(i)$ based on similarity-weighted probability $\propto 4s_{ij}(1 - s_{ij})$, following Axelrod's principle that moderate similarity maximizes interaction likelihood.
2. Agent i sends its complete cultural vector $\mathbf{T}_i$ to j.
3. Agent j evaluates whether to adopt a trait from i through LLM-based reasoning (detailed below).
4. If adoption occurs, agent j updates one differing trait dimension to match i's value.

This process repeats until convergence or maximum timesteps ($T_{\text{max}} = 50$).

Key parameters. Unlike traditional models where adoption is deterministic or purely similarity-based, we introduce openness $\alpha \in \{\text{low}, \text{medium}, \text{high}\}$ as an independent psychological parameter modulating receptivity. Combined with information flow $k \in \{1, 3, 5\}$, this yields a 3×3 parameter space enabling systematic exploration of psychological-structural interactions.

3.2 LLM AGENT ARCHITECTURE

Model selection. We employ Qwen3-8B (Yang et al., 2025), a state-of-the-art open-source language model with 8 billion parameters trained on diverse multilingual data. Qwen3 demonstrates strong performance on reasoning tasks, instruction following, and contextual understanding—capabilities essential for cultural reasoning. We chose this model over alternatives (GPT-4, Claude, Llama) to ensure reproducibility through open access while maintaining sufficient sophistication for complex decision-making.

Agent instantiation. Each agent is a separate instance of the LLM with consistent personality configuration. Unlike shared-model approaches where one LLM simulates all agents sequentially, our design ensures parallelism and independence: each agent maintains its own conversation history, personality traits, and decision context. This architecture prevents information leakage between agents and enables concurrent processing.

Decision mechanism. Traditional agents adopt traits through stochastic rules:

$$P_{\text{adopt}}^{\text{trad}}(i,j) = s_{ij} \cdot \alpha_i + \epsilon \tag{3}$$

where α_i is a fixed openness parameter and ϵ represents random noise. This approach is tractable but psychologically unrealistic—humans don't mechanically compute similarity-weighted probabilities.

Our LLM agents instead engage in contextual reasoning:

$$P_{\text{adopt}}^{\text{LLM}}(i,j) = \text{LLM}(\alpha_i, s_{ij}, \mathbf{T}_i, \mathbf{T}_j, H_i, \text{social context}) \tag{4}$$

where H_i represents interaction history and social context includes information about multiple neighbors. The LLM evaluates trait compatibility, considers personal preferences, weighs social influence, and makes reasoned decisions expressed through natural language before selecting actions.

Prompt engineering. Agent behavior is controlled through carefully designed prompts consisting of three components:

System prompt establishes the simulation context and cultural dynamics principles:

```
You are a CulturalAgent in a simulation based on Axelrod's
cultural dissemination model. You have cultural traits
across five dimensions: music preference, culinary
preference, fashion style, political orientation, leisure
activity.
```

Task instruction specifies the decision problem. For target selection:

```
Choose ONE agent from your social network to share your
cultural traits with. Select a target and explain why
you chose them. Return your decision in JSON format:
{"target_id": "<ID>", "reason": "<brief explanation in
1-2 sentences>"}
OBSERVATION: YOUR CULTURAL PROFILE: music preference: jazz,
culinary preference: mediterranean, ... POTENTIAL TARGETS:
Agent 001, Similarity: 0.60, Traits: music: classical,
...
```

For trait adoption:

```
You've received cultural recommendations from another agent.
From the differences between your cultures, choose ONE
dimension to adopt from them. Consider: (1) How well this
trait aligns with your existing cultural identity. (2) The
influence of your personality (openness level). (3) The
recommender's cultural coherence. Return your decision
in JSON format: {"adopt_dimension": "<dimension_name>",
"reasoning": "<brief explanation>"}
OBSERVATION: Recommender: Agent 001, Similarity: 0.60,
Your personality: high openness. CULTURAL DIFFERENCES:
music preference - Your current: jazz, Their value:
classical; ...
```

Openness operationalization. The openness parameter $\alpha \in \{\text{low}, \text{medium}, \text{high}\}$, is implemented through the injection of personality-specific cues into the observation prompt:

- Low: "Your personality: conservative and stability-seeking. You prefer maintaining your existing cultural traits and only adopt new traits when there is very high similarity with others. Change makes you uncomfortable, and you value cultural consistency."
- Medium: "Your personality: moderately open to change. You consider adopting traits when there is reasonable cultural overlap and the trait seems valuable. You balance maintaining your identity with selective cultural adaptation."
- High: "Your personality: highly open to new experiences. You actively seek cultural diversity and readily consider adopting traits from others even with moderate cultural overlap. You view cultural exchange as enriching and embrace change."

Importantly, we observed qualitative differences in reasoning patterns. Low openness agents frequently cited "maintaining cultural identity" and "preference for stability." Medium openness agents balanced "opportunity for enrichment" with "consistency concerns." High openness agents emphasized "learning from diversity" and "cultural exploration." These emergent reasoning patterns were not explicitly scripted, suggesting the LLM successfully internalized personality constructs.

3.3 Experimental Design

Factorial structure. We conduct a full 3×3 factorial experiment crossing three openness levels ($\alpha \in \{\text{low}, \text{medium}, \text{high}\}$) with three information flow orders ($k \in \{1, 3, 5\}$). This yields nine experimental conditions. Each condition is replicated three times with different random seeds ($\{42, 123, 456\}$) to control for stochastic variation in both cultural trait initialization and LLM sampling. Total: $N = 27$ independent simulation runs.

Population and topology. Each simulation instantiates 100 agents ($N = 100$) arranged on a 10×10 square lattice with periodic boundary conditions (toroidal topology eliminating edge effects). This population size balances computational feasibility with adequate statistical power for detecting interaction effects.

Initialization. Cultural traits are initialized randomly with uniform distribution: for each agent i and dimension d, $t_{id} \sim \text{Uniform}(\{0, \ldots, 9\})$ independently. This maximizes initial cultural diversity ($\mathbb{E}[s_{ij}] = 0.1$ for random agent pairs), creating a challenging scenario for convergence and ensuring that observed patterns emerge from dynamics rather than initial conditions.

Simulation protocol. Each simulation runs for $T_{\text{max}} = 50$ timesteps. At each step, agents undergo four sequential phases. In the *target selection* phase, every agent i simultaneously selects an interaction partner $j \in N_k(i)$ via the LLM-based target selection procedure. The probability of choosing j is similarity-weighted, $\Pr(i \to j) \propto 4s_{ij}(1 - s_{ij})$, following the Axelrod interaction function that peaks at $s = 0.5$. During the *trait transmission* phase, if $s_{ij} \in (0, 1)$, agent j decides whether to adopt one of i's traits using the LLM-based adoption procedure; no transmission occurs when $s_{ij} = 0$ or 1. In the *update* phase, adopting agents modify one differing trait dimension to match the recommender's value, with the specific dimension selected by the LLM. Finally, in the *monitoring* phase, the Cultural Homogeneity Index (CHI) is computed after each step. Simulations typically stabilize (CHI variance < 0.01 over 10 consecutive steps) between steps 35–45.

Computational resources. Simulations were executed on NVIDIA A100 GPUs (40GB) using PyTorch 2.0 and the `transformers` library (v4.35). Each run required approximately 2.5 hours wall-clock time, totaling $\sim$67.5 GPU-hours across all conditions.

LLM configuration. Qwen3-8B was used with temperature $\{0.5, 0.7, 0.9\}$ (depending on openness level), top-p=0.9, max tokens=4096, presence and frequency penalties=0.0, and batch size=100 for parallel agent processing.

3.4 Evaluation Metrics

Cultural Homogeneity Index (CHI). Our primary outcome measure quantifies population-level cultural convergence:

$$\text{CHI}(t) = \frac{1}{D}\sum_{d=1}^{D}\max_{v\in\mathcal{V}}\frac{|\{i : t_{id}(t) = v\}|}{N} \tag{5}$$

For each dimension d, we compute the proportion of agents adopting the most common trait value, then average across dimensions. CHI ranges from $\frac{1}{q} = 0.1$ (perfect diversity, each value equally represented) to 1.0 (perfect homogeneity, all agents share the most common trait in every dimension).

This metric captures dimension-wise dominance rather than requiring complete cultural identity across agents. A population where 70% of agents share the same music preference, 60% share culinary preference, etc., would have higher CHI than one with uniform distributions, even if no two agents have identical complete cultural vectors.

4 Results

Our experiments reveal that individual openness and information flow jointly shape cultural dynamics through non-additive mechanisms. Analysis of 27 simulation runs across a 3×3 factorial design establishes three key findings: (1) openness exhibits nonlinear effects with cascade amplification at moderate levels, (2) information flow demonstrates a robust 3rd-order optimum contradicting monotonic connectivity assumptions, and (3) their interaction produces capacity-connectivity mismatches where excessive networks impair conservative agents.

4.1 Openness Effects: Cascade Amplification Regime

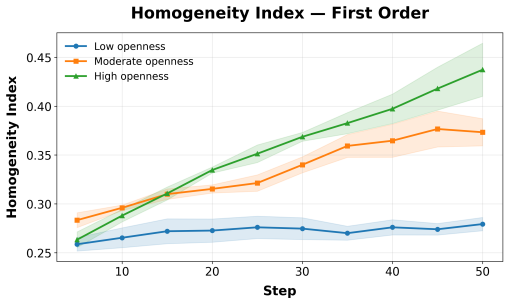
Statistical analysis confirms strong openness effects. Fractional logit regression yields $\beta = 0.305$ ($z = 7.59$, $p < 0.001$), with median CHI values of 0.279-0.284 (low), 0.373-0.427 (moderate), and 0.433-0.489 (high) depending on interaction order. Spearman rank correlation demonstrates monotonic relationship ($\rho = 0.896$, $p = 0.001$).

Figure 2a displays temporal trajectories under 1st-order interactions. High openness achieves CHI = 0.437 versus 0.279 for low openness—a 57% advantage emerging after step 15 when high openness accelerates (slope $\approx$ 0.0035 per step) while low openness plateaus. Crucially, the relationship between openness and convergence is not uniformly linear across network topologies. Moving from low to moderate openness produces asymmetric gains: ΔCHI = 0.094 for 1st order (34%), 0.157 for 3rd order (55%), and 0.102 for 5th order (31%). This asymmetry reveals that moderate openness achieves disproportionate benefits specifically under optimal network conditions (3rd order), rather than uniform scaling—suggesting an interaction between psychological receptivity and structural affordances that we explore in Section 4.3.

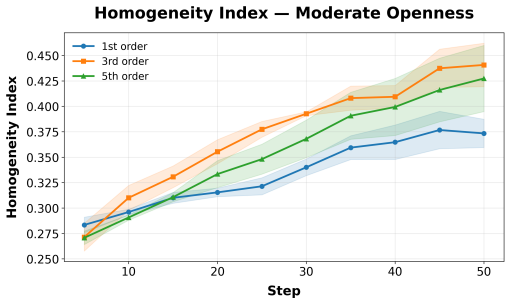
Trajectory signatures encode distinct mechanisms. Beyond final convergence levels, the *shape* of temporal trajectories reveals underlying dynamics. Low openness ($\approx$0.28) exhibits near-linear shallow growth with total change ΔCHI $\approx$ 0.012 over 50 steps—consistent with frozen cultural landscapes where agents resist both initial nucleation and long-term refinement. This stagnation occurs because conservative agents rarely adopt traits even from highly similar neighbors, preventing the formation of dominant cultural patterns.

High openness shows front-loaded convergence: rapid early growth (steps 5-20, slope $\approx$0.006 per step) transitioning to gradual saturation by step 40. This pattern reflects high receptivity enabling quick alignment with neighbors, but convergence slows as remaining heterogeneity becomes entrenched in spatially separated regions that require long-range bridges to resolve.

Most revealing, moderate openness (0.373) displays *accelerating* convergence where the slope *increases* between steps 20-40 (from $\approx$0.002 to $\approx$0.004 per step) before leveling off. This acceleration is theoretically significant: in systems with independent stochastic adoption processes, convergence rates should remain constant or decelerate as cultural variation depletes and adoption opportunities diminish. Acceleration necessarily implies positive feedback dynamics—early convergence must somehow facilitate subsequent convergence.



(a) **Temporal dynamics reveal nonlinear openness effects and 3rd-order optimization.** Under 1st-order interactions, high openness (green, 0.437) achieves 57% advantage over low openness (blue, 0.279), with moderate openness (orange, 0.373) showing mid-trajectory acceleration indicative of cascade dynamics.



(b) **Temporal dynamics reveal nonlinear openness effects and 3rd-order optimization.** For moderate openness, 3rd-order (orange, 0.441) outperforms both 1st-order (blue, 0.373, +18%) and 5th-order (green, 0.427, +14% vs. 1st), with divergence emerging post-step 25.

Cascade amplification hypothesis. We propose that acceleration at moderate openness arises from self-reinforcing cultural cascades. As agents selectively adopt traits from similar neighbors, this creates emergent local cultural clusters (e.g., spatial regions where 60-70% share a trait value). These clusters amplify social influence: when an agent interacts with a cluster member, the recommendation now carries implicit endorsement from the broader cluster, increasing adoption likelihood. This positive feedback enables cascades where local convergence triggers broader convergence.

Why does this occur specifically at moderate openness? Low openness agents resist social amplification—even strong cluster signals fail to overcome their conservative threshold, preventing cascade propagation. High openness agents converge so rapidly that the population reaches near-homogeneity before spatial clusters can form and propagate—cascades are pre-empted by direct convergence. Moderate openness occupies a critical regime where: (1) initial adoptions are selective enough to create coherent local clusters rather than random noise, (2) adoption frequency suffices to propagate these clusters through spatial neighborhoods, (3) convergence proceeds slowly enough that cascade dynamics can develop and achieve population-level impact before saturation. This cascade-susceptible regime implies that threshold-targeted interventions focused on moderately receptive populations can maximize policy leverage by triggering self-reinforcing dynamics.

4.2 Information Flow Effects: The 3rd-Order Optimum

Figure 2b reveals a striking non-monotonic pattern where 3rd-order interactions achieve maximal convergence across all openness levels. For moderate openness: 1st-order reaches 0.373, 3rd-order achieves 0.441 (+18%), while 5th-order regresses to 0.427. The heatmap (Figure 3a) demonstrates universal applicability: high openness shows 1st = 0.437, **3rd = 0.489** (highest overall), 5th = 0.433; moderate shows 1st = 0.373, **3rd = 0.441**, 5th = 0.427; low shows 1st = 0.279, 3rd = 0.284, 5th = 0.325. Notably, 1st and 5th-order in high openness converge to nearly identical levels (0.437 vs. 0.433) despite vastly different topologies (4 vs. 60 neighbors), suggesting both insufficient and excessive connectivity similarly constrain transmission through different mechanisms.

Exploration-exploitation trade-off. We hypothesize 3rd-order ($\approx$28 neighbors) optimally balances competing forces. *Exploration* requires sufficient connectivity to escape local equilibria—1st-order creates spatial bottlenecks trapping populations when regions develop suboptimal configurations. *Exploitation* requires signal coherence—5th-order's 60 neighbors exhibit high diversity because they span large spatial scales without local homogenization. When agents receive recommendations from 60 culturally diverse sources, signals become contradictory, diluting any single attractor and preventing coordinated consensus formation. At 3rd-order, agents access sufficient diversity to escape traps while maintaining spatial coherence where neighbors-of-neighbors share enough traits to generate mutually reinforcing signals.

Two-phase dynamics. Temporal trajectories reveal connectivity operates differently across evolution. During the *nucleation phase* (steps 5-25), all orders show similar slopes—random initialization ensures even local neighbors provide diverse exposure, rendering extended connectivity redun-

dant. The *consolidation phase* (steps 25-50) shows sharp divergence: 3rd-order maintains growth ($\Delta H/\Delta t \approx 0.0027$ per step, 0.392 $\rightarrow$ 0.441), while 1st-order plateaus (stagnating at 0.373) as agents become trapped within homogeneous local neighborhoods. The 5th-order trajectory tracks 3rd-order until step 30, then exhibits reduced rates, suggesting noise saturation onset. This phase transition—connectivity irrelevant early, critical late—implies network expansion investments should target post-clustering dynamics (timestep 20-25+) rather than initial convergence.

Low-openness exception. Uniquely, low openness shows monotonic improvement (0.279 $\rightarrow$ 0.284 $\rightarrow$ 0.325), revealing the 3rd-order optimum depends on population psychology. Conservative agents benefit from raw exposure rather than signal coherence—their selective threshold means they never experience noise overload. However, even 5th-order achieves only 0.325, representing 40% below high-3rd (0.489), demonstrating structural interventions cannot overcome psychological barriers.

4.3 Interaction Effects: Capacity-Connectivity Matching

Two-way ANOVA reveals significant interaction ($F(4, 18) = 3.45$, $p = 0.028$, $\eta_p^2 = 0.43$) with effect size comparable to flow main effect ($\eta_p^2 = 0.49$), indicating connectivity effects depend on openness. This contradicts additive models where openness and flow independently contribute.

Figure 3a visualizes the 75% performance gap between low-1st (0.279) and high-3rd (0.489), reflecting synergistic effects. For high openness, extended flow produces non-monotonic outcomes: 0.437 (1st) $\rightarrow$ **0.489 (3rd, +12%)** $\rightarrow$ 0.433 (5th, -11% vs. peak), contradicting assumptions that broader networks monotonically enhance transmission. Low openness shows monotonic improvement (0.279 $\rightarrow$ 0.325) but remains 40% below peak even at maximal connectivity, revealing fundamental ceiling effects where structural interventions hit diminishing returns against psychological constraints.

Capacity-connectivity matching principle. The non-monotonic pattern emerges from mismatch between *adoption capacity* (psychological receptivity determining adoption frequency) and *information bandwidth* (network connectivity determining available influences). High openness under 1st-order faces *exposure bottlenecks*—agents exhaust local variation, achieving local consensus but lacking bridges to other regions. Under 3rd-order, sufficient diversity enables exploration while maintaining spatial coherence for exploitation—agents effectively integrate signals across 28 influences, extracting dominant patterns. Under 5th-order, *noise saturation* emerges—60 culturally scattered sources provide contradictory signals, preventing coordination. Adoptions scatter across 60 sources rather than reinforcing single attractors, transforming cultural evolution from directional convergence into random walk.

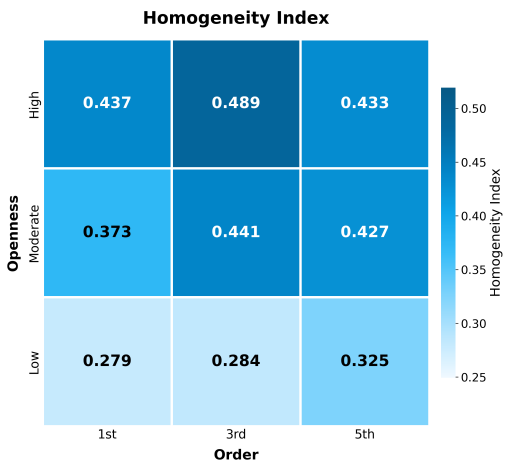
Low openness agents adopt so infrequently that the bottleneck is encountering *any* influence strong enough to breach their adoption threshold, not signal-to-noise ratio. More neighbors increase chances of finding highly similar agents, hence monotonic improvement. However, absolute ceiling (0.325) demonstrates structural interventions cannot compensate for psychological barriers.

Implications. The matching principle reveals optimal network architecture must adapt to population characteristics. Uniform global connectivity (5th-order) benefits conservative users but harms open users through noise saturation. Localized connectivity (1st-order) protects open users from noise but starves conservative users. The 3rd-order regime ($\approx$20-30 neighbors) provides robust performance across moderate-to-high openness levels, suggesting this as a reasonable default for heterogeneous populations, though truly optimal design might require adaptive networks that increase connectivity for low-openness segments while maintaining intermediate levels for high-openness segments.

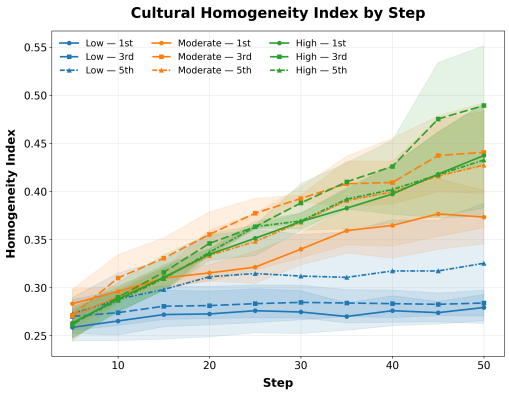
5 Discussion

Theoretical implications. Our findings support the framework that cultural dynamics emerge from psychological-structural interplay. The significant main effects demonstrate that openness and flow operate as independent mechanisms. The interaction reveals non-additive effects: optimal network structure depends on population characteristics. This suggests interventions must jointly consider individual attitudes and communication infrastructure.

Practical implications. For system designers: (1) Conservative populations may maintain stability through localized interactions; (2) Open populations benefit from global connectivity enabling cross-cultural exchange; (3) One-size-fits-all network policies may fail—structure should match population



(a) **Interaction effects reveal capacity-connectivity matching.** Heatmap shows 3rd-order achieves peak convergence across all openness levels (0.489 high, 0.441 moderate, 0.284 low), demonstrating robust inverted-U relationship. The 75% gap between low-1st (0.279) and high-3rd (0.489) reflects synergistic effects.



(b) **Interaction effects reveal capacity-connectivity matching.** Temporal trajectories: 3rd-order (solid lines) maintains steepest late-phase slopes for moderate/high openness, while 5th-order (dashed) exhibits early tracking followed by saturation due to noise overload. Low openness shows minimal order separation, confirming psychological ceilings dominate.

characteristics. For policymakers: integration efforts should assess community openness before prescribing connectivity interventions.

Limitations. Our model simplifies reality in several ways. Grid topology ignores small-world and scale-free properties. Discrete traits miss continuous cultural dimensions. LLM agents, while more sophisticated than rule-based counterparts, still reflect training data biases and may not fully capture human irrationality, emotion, and context-dependent behavior. Computational constraints limit population size (100 agents) and duration (50 steps) compared to real societies. Static networks ignore dynamic coevolution of structure and culture.

Future work would address limitations by: (1) testing on realistic network topologies; (2) extending to continuous cultural spaces; (3) validating against empirical diffusion data; (4) investigating LLM bias effects through multi-model comparisons; (5) scaling to larger populations via hybrid approaches.

Broader impacts. Positive applications include informing integration policies, platform design for constructive dialogue, and organizational culture management. Risks include manipulation for political control, privacy violations through cultural surveillance, and inadvertent suppression of diversity. While our work involves only artificial agents, real-world deployment requires ethical safeguards, informed consent, transparency, and multi-stakeholder governance.

6 Conclusion

This research demonstrates that individual openness and information flow jointly determine cultural fragmentation in LLM-based multi-agent systems through independent but synergistic mechanisms. Using Qwen3-8B agents in a 3×3 experimental design, we provide quantitative evidence that higher openness and expanded information flow both significantly reduce fragmentation, with optimal outcomes achieved through their combination. The key contribution lies in decoupling psychological and structural factors using cognitively sophisticated AI agents, revealing that effective interventions should target both dimensions simultaneously—promoting individual openness and optimizing communication ranges. Future research should extend this framework to realistic network topologies, dynamic parameters, and empirical validation. This computational approach provides a methodological foundation for advancing quantitative understanding of cultural dynamics in both artificial and natural social systems.

REFERENCES

Robert Axelrod. The dissemination of culture: A model with local convergence and global polarization. *Journal of conflict resolution*, 41(2):203–226, 1997.

Lauro A. Barbosa and José F. Fontanari. Culture-area relation in axelrod's model for culture dissemination, 2009. URL https://arxiv.org/abs/0902.1346.

C. Gracia-Lazaro, L. F. Lafuerza, L. M. Floria, and Y. Moreno. Residential segregation and cultural dissemination: An axelrod-schelling model, 2009. URL https://arxiv.org/abs/0907.2360.

Carlos Gracia-Lázaro, Edgardo Brigatti, Alexis R Hernández, and Yamir Moreno. Polarization inhibits the phase transition of axelrod's model. *Physical Review E*, 103(6):062306, 2021.

Pádraig MacCarron, Paul J Maher, Susan Fennell, Kevin Burke, James P Gleeson, Kevin Durrheim, and Michael Quayle. Agreement threshold on axelrod's model of cultural dissemination. *Plos one*, 15(6):e0233995, 2020.

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.

Lucía Pedraza, Sebastián Pinto, Juan Pablo Pinasco, and Pablo Balenzuela. A novel analytical formulation of the axelrod model, 2020. URL https://arxiv.org/abs/2006.15241.

Lucas R. Peres and José F. Fontanari. Mass media destabilizes the cultural homogeneous regime in axelrod's model, 2009. URL https://arxiv.org/abs/0910.0866.

Sebastián Pinto and Pablo Balenzuela. Erdös-rényi phase transition in the axelrod model on complete graphs, 2020. URL https://arxiv.org/abs/1912.09420.

Arezky H Rodríguez and Y Moreno. Effects of mass media action on the axelrod model with social influence. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, 82(1):016111, 2010.

Alex Stivala, Garry Robins, Yoshihisa Kashima, and Michael Kirley. Ultrametric distribution of culture vectors in an extended axelrod model of cultural dissemination. *Scientific reports*, 4(1): 4870, 2014.

Lei Wang, Chen Ma, Xueyang Feng, Zeyu Zhang, Hao Yang, Jingsen Zhang, Zhiyuan Chen, Jiakai Tang, Xu Chen, Yankai Lin, et al. A survey on large language model based autonomous agents. *Frontiers of Computer Science*, 18(6):186345, 2024.

Ruoxi Xu, Yingfei Sun, Mengjie Ren, Shiguang Guo, Ruotong Pan, Hongyu Lin, Le Sun, and Xianpei Han. Ai for social science and social science of ai: A survey, 2024. URL https://arxiv.org/abs/2401.11839.

An Yang, Anfeng Li, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chang Gao, Chengen Huang, Chenxu Lv, et al. Qwen3 technical report. *arXiv preprint arXiv:2505.09388*, 2025.

A COMPUTATIONAL RESOURCES

All experiments were conducted on NVIDIA A100 GPUs with 40GB memory using PyTorch 2.0 and transformers library version 4.35.0. Each simulation required approximately 2-3 hours of computation time depending on the convergence rate. Each experiment was replicated three times across conditions to ensure reproducibility while maintaining statistical independence.

LLM Configuration: Qwen3-8B was configured with temperature=0.7, top-p=0.9, max_tokens=4096, and presence_penalty=0.0 to balance reasoning consistency with behavioral variability.