A Spectral Theory of Memetic Propagation:

Understanding Idea Spread Through Frequency and Resonance

Jonas Hallgren

February 27, 2026

Abstract

Why do some ideas spread universally while others fragment communities? We introduce a mathematical framework that describes memes as messages carrying spectral signatures—frequency content that determines both their spreadability and their impact on agent behavior. Low-frequency messages propose gentle, diffuse shifts in action policies and spread easily because they induce small cognitive dissonance. High-frequency messages demand sharp, specific behavioral reorganization and spread slowly despite capturing attention. Intermediate-frequency messages occupy a sweet spot: structured enough to be meaningful, smooth enough to bridge communities. By viewing ideas through both geometric (spectral patterns on semantic graphs) and information-theoretic (probability distributions over action policies) lenses, we unify intuitions about resonance, cognitive dissonance, and collective discourse dynamics. This framework makes testable predictions about polarization, bridging, and the architecture of sustainable dialogue.

1 Introduction: The Puzzle of Differential Spread

Consider three messages about climate and collective action:

A. *"We should care for future generations"*

B. *"Climate change requires coordinated international action"*

C. *"Implement a $50/ton carbon tax starting January 2026, with agricultural exemptions through 2028"*

If you posted these sequentially to politically diverse forums—spanning environmental activists, libertarian economists, rural farmers, urban progressives—something remarkable would emerge in the response patterns. The first message would resonate almost universally, generating broad agreement with minimal controversy. The third would create sharp boundaries: clusters of strong support, clusters of strong opposition, with little middle ground. The second would occupy intermediate territory: bridging some communities while dividing others, creating structured patterns of partial agreement.

This differential spread reveals something fundamental about how ideas propagate through social networks. It's not random variation or mere differences in content complexity. Rather, these messages carry different *spectral signatures*—patterns that determine how they resonate with and transform the belief states of agents who encounter them.

1.1 Three Views of the Same Phenomenon

The power of spectral thinking emerges when we recognize that what appears as three separate observations—geometric patterns on graphs, probability distributions over beliefs, behavioral policy updates—actually represent three perspectives on a unified underlying structure. Figure **??** illustrates this correspondence for our climate messages.

The Geometric View: When we visualize the semantic landscape—imagine a graph where nodes represent positions in ideological or conceptual space, and edges connect similar positions—each message creates a pattern. Message A colors this graph almost uniformly: ask people across the political spectrum whether they care about future generations, and you find smooth gradients of agreement. Message C creates sharp discontinuities: support clusters in specific regions (say, environmental economists and climate activists) while generating strong opposition elsewhere (libertarians concerned about government overreach, industry representatives worried about competitiveness). The smoothness or sharpness of these patterns is precisely what spectral graph theory captures through eigenvector decomposition.

The Distributional View: Each message implicitly proposes a probability distribution over action policies—not abstract beliefs, but concrete behavioral implications. "Care for future generations" is diffuse and low-amplitude: it gently nudges your priorities across many possible actions without demanding any specific course. "$50/ton carbon tax starting January 2026" is peaked and high-amplitude: accepting this message concentrates your probability mass on very specific political actions—how you vote, which organizations you support, what policies you advocate for. The information-theoretic distance between an agent's current distribution and the message's proposed distribution, measured through KL divergence, quantifies the cognitive dissonance of adoption.

The Behavioral View: What makes these different frequency contents consequential is their connection to action. A low-frequency message asks for small updates to your existing behavioral patterns—perhaps you donate slightly more to long-term causes, or think a bit more carefully about sustainability, but your day-to-day life proceeds largely unchanged. A high-frequency message demands wholesale reorganization: if you genuinely accept that specific policy prescription, you must shift your voting behavior, your public advocacy, your resource allocation, potentially your career focus. The amplitude of the distribution over action policies captures precisely this demanded magnitude of behavioral change.

These three views are not separate claims requiring independent validation. They are different languages for expressing the same underlying phenomenon: *messages have spectral structure that determines how they propagate through networks of agents.* A smooth pattern on a semantic graph *is* a diffuse distribution over policies *is* a small required behavioral update. A sharp pattern *is* a peaked distribution *is* a large required reorganization.

1.2 Cool graph

1.3 Why Conspiracy Theories Resist Spread

This unified perspective illuminates puzzles in discourse dynamics. Consider conspiracy theories—claims like "the moon landing was faked" or "vaccines contain tracking microchips." These messages often exhibit a paradox: they generate intense engagement and discussion (suggesting high salience) yet fail to spread beyond specific communities (suggesting low fitness).

The spectral view resolves this. Conspiracy theories are profoundly high-frequency messages. They propose not just factual claims but entire reorganizations of epistemic policy: *whom to trust for information, which institutions to engage with, how to educate your children, what medical decisions to make.* The probability distribution over action policies they propose is both high-

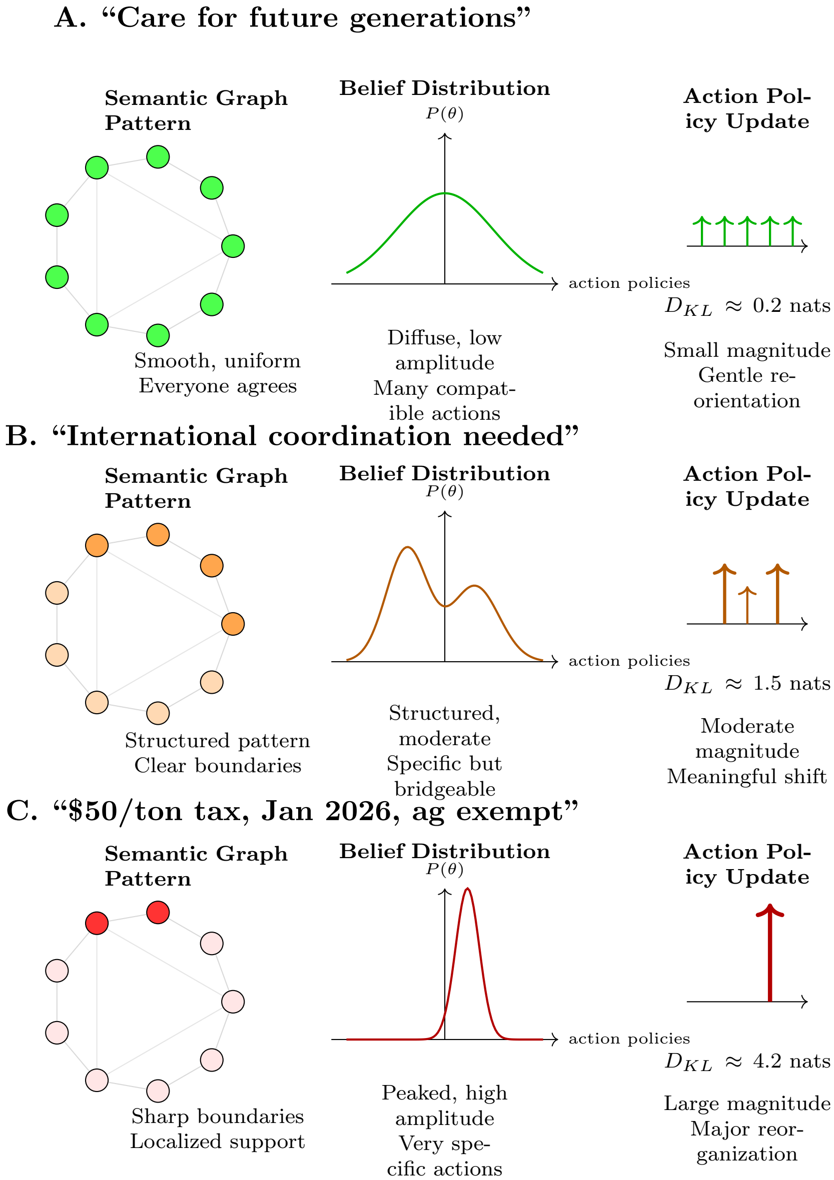

Figure 1: Three perspectives on message frequency content. Each message can be understood through (1) its geometric pattern on the semantic graph, (2) its implied probability distribution over action policies, and (3) the magnitude of behavioral update it demands. **Low-frequency messages** create smooth agreement patterns, propose diffuse low-amplitude distributions, and require only gentle reorientation—spreading easily because they induce small cognitive dissonance (D_{KL}). **High-frequency messages** create sharp community boundaries, propose peaked high-amplitude distributions over specific actions, and demand major behavioral reorganization—spreading slowly despite high salience. **Intermediate-frequency messages** exhibit structured patterns, propose moderately peaked distributions, and require meaningful but not extreme updates—occupying the sweet spot for sustained discourse.

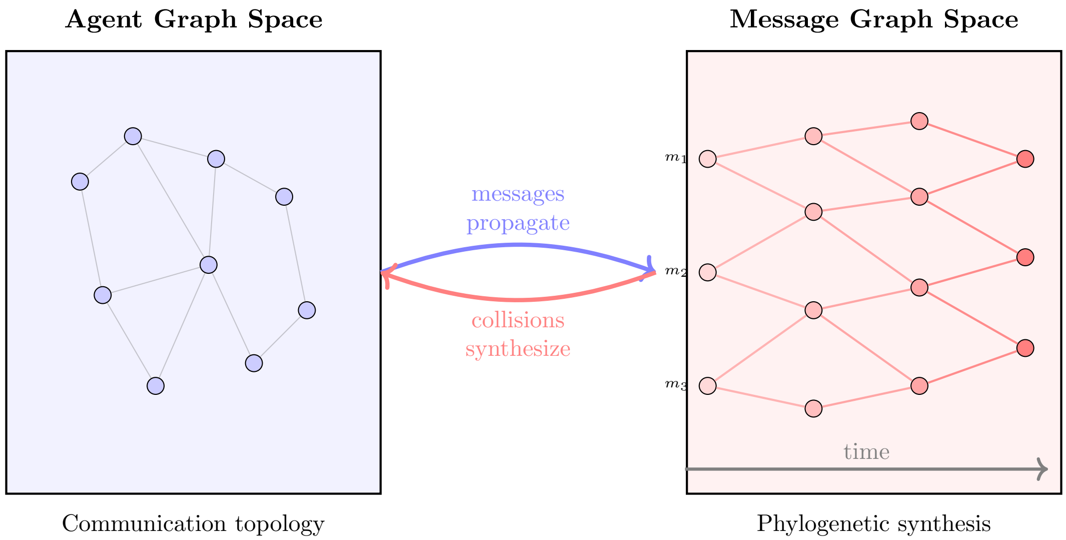

Figure 2: The dual-space architecture of memetic evolution. **Left**: The agent graph represents communication topology through which messages flow. **Right**: The message space exhibits phylogenetic structure with time flowing left to right. Ancestral messages (lighter nodes, left margin) give rise through collision and synthesis to descendant messages (darker nodes, progressing rightward). The reticulated branching pattern—where multiple parent messages converge to produce single offspring—captures the distinctive feature of cultural evolution: unlike biological phylogeny where lineages strictly diverge, memetic lineages can reconverge through synthesis when ideas from different families collide in agent minds. This tree structure reveals cultural innovation as an evolutionary process of recombination, not mere transmission.

amplitude (demanding major behavioral changes) and peaked in regions far from most agents' current positions.

For someone currently trusting mainstream institutions, adopting a conspiracy theory requires relocating vast amounts of probability mass across action policy space—an update with KL divergence of many nats. This creates fierce resistance: the message induces such large cognitive dissonance that exponentially few agents adopt it. Yet within communities that have already shifted toward those peaked distributions, the message spreads easily—it requires only small additional updates from an already-aligned prior.

This explains the characteristic pattern: conspiracy theories show high within-community spread (where they're effectively low-frequency relative to local beliefs) but fail to bridge communities (where they appear as extreme high-frequency perturbations). The spectral signature is intrinsic to the message's content—the behavioral reorganization it demands—but its effective frequency varies with the local belief landscape.

1.4 The Intermediate Frequency Phenomenon

The most intriguing prediction of spectral thinking concerns intermediate frequencies. Pure low-frequency messages, while spreadable, lack distinctiveness—they generate little engagement precisely because they demand little change. Pure high-frequency messages, while attention-grabbing, resist adoption beyond committed communities. But intermediate-frequency messages—those proposing structured, moderately peaked distributions over action policies—can be both salient and spreadable.

Message B exemplifies this. "Climate change requires coordinated international action" makes a meaningful claim that structures behavior: it suggests supporting international agreements, engaging with global governance questions, thinking beyond national boundaries. The distribution over action policies has clear structure—not everything is permitted—but isn't maximally peaked either. Multiple paths forward remain compatible: carbon pricing, technology transfer, adaptation funding, regulatory harmonization. The moderate amplitude allows bridging between communities that share some low-frequency alignment (caring about futures, valuing cooperation) while providing enough structure to be actionable and interesting.

This creates the characteristic signature of productive discourse: intermediate-frequency messages dominate sustained conversations because they optimize the tradeoff between spreadability (determined by cognitive dissonance) and salience (determined by distinctiveness from background). They're smooth enough to cross community boundaries, sharp enough to matter. They induce moderate KL divergence—large enough to be informative, small enough to be adoptable.

1.5 Contributions and Structure

This paper develops the mathematical foundations of spectral memetics—understanding message propagation through frequency content. We show that:

1. Messages naturally decompose into frequency components via graph Laplacian eigenvectors, with frequency content determining both geometric patterns (smooth vs. sharp on semantic graphs) and distributional properties (diffuse vs. peaked over action policies)

2. The spread rate of messages decreases with frequency content because high-frequency messages induce large cognitive dissonance (KL divergence), creating exponential resistance to adoption

3. Equilibrium discourse exhibits intermediate-frequency dominance: despite low-frequency messages being most spreadable and high-frequency messages being most salient, intermediate frequencies capture disproportionate sustained engagement

4. Polarization corresponds to loss of low-frequency overlap between communities—when groups lose shared smooth structure over action policies, even gentle messages induce large mutual KL divergence

5. Effective bridging operates at intermediate frequencies where messages have enough structure to matter but enough smoothness to cross community boundaries

Section **??** develops the mathematical framework: spectral decomposition of semantic space, agent belief dynamics as probability distributions over action policies, and message propagation governed by frequency-dependent cognitive dissonance. Section **??** analyzes the resulting dynamical system, revealing phase transitions, equilibrium structures, and the emergence of frequency-dependent selection. Section **??** details experimental protocols for validating these predictions using discourse data. Section **??** connects the framework to broader questions in collective intelligence and identifies future research directions.

2 Mathematical Foundations

2.1 The Dual-Space Architecture

Our framework rests on representing two distinct but coupled aspects of memetic propagation:

Definition 1 (Agent Graph)**.** *The* agent graph $G_A = (V, E_A)$ *represents entities capable of holding and updating beliefs: humans, AI systems, organizations. Vertices* $v \in V$ *are agents; edges* $(i, j) \in E_A$ *represent communication pathways through which messages propagate.*

Definition 2 (Message Graph)**.** *The* message graph $G_M = (M, E_M)$ *represents semantic relationships between messages. A message* $m \in M$ *is a discrete communicative unit carrying implicit or explicit claims about the world. Edges* $(m_\alpha, m_\beta) \in E_M$ *represent semantic similarity: messages are connected if they concern related topics, support similar policies, or invoke related concepts.*

The architecture is *dual* because these graphs have autonomous properties but are coupled through message-agent interactions. Messages propagate through the agent substrate (flows in G_A) while their semantic relationships (structure of G_M) determine their frequency content. This separation proves analytically crucial: G_A evolves slowly (social connections change on timescales of months) while message dynamics unfold rapidly (hours to days), enabling quasi-static approximations.

2.2 Spectral Decomposition of Semantic Space

The message graph induces a natural frequency basis through its Laplacian operator. Let A_M denote the adjacency matrix (with $A_M(m_\alpha, m_\beta)$ reflecting semantic similarity) and D_M the degree matrix. The *graph Laplacian* is:

$$L_M = D_M - A_M \tag{1}$$

This operator measures how messages vary across semantic space: for any function $f : M \to \mathbb{R}$ assigning values to messages, $L_M f$ captures the discrete derivative—how rapidly f changes between semantically similar messages.

The Laplacian's spectral decomposition provides the frequency basis:

$$L_M \phi_k = \lambda_k \phi_k \tag{2}$$

where $\{\phi_k\}_{k=0}^{|M|-1}$ are orthonormal eigenvectors and $\{0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \cdots\}$ are eigenvalues. This decomposition has immediate geometric interpretation:

- ϕ_0 is constant across all messages (the trivial mode representing uniform agreement)
- Low-index eigenvectors ϕ_k with small λ_k vary smoothly: $\phi_k(m_\alpha) \approx \phi_k(m_\beta)$ when messages are semantically similar
- High-index eigenvectors ϕ_k with large λ_k vary sharply: ϕ_k oscillates rapidly across semantic neighborhoods

Any pattern over messages decomposes uniquely into these frequency components:

$$f = \sum_{k=0}^{|M|-1} c_k \phi_k \tag{3}$$

where coefficients $c_k = \langle f, \phi_k \rangle$ measure how much pattern f projects onto frequency mode k.

Definition 3 (Message Frequency Content)**.** *A message* m *has* low-frequency content *if its semantic influence (how it affects nearby messages in the graph) is captured primarily by low-eigenvalue eigenvectors—it creates smooth patterns of agreement. It has* high-frequency content *if it projects primarily onto high-eigenvalue eigenvectors—it creates sharp boundaries and localized responses.*

The mathematical formalism captures our intuitive examples. "Care for future generations" creates nearly uniform coloring across the semantic graph (high projection onto ϕ_0, minimal higher frequency components). "Specific carbon tax policy" creates localized peaks of support with sharp boundaries (minimal low-frequency content, high projection onto large-λ_k eigenvectors).

2.3 Action Policy Distributions and Cognitive Dissonance

Each agent $i \in V$ maintains a belief state b_i, but crucially, we represent this as a *probability distribution over action policies* rather than abstract propositions. Let Θ denote the space of possible behavioral policies—concrete specifications of how an agent acts in various circumstances. An agent's belief state $P_i(\theta)$ captures their uncertainty and commitments over Θ.

This action-centric representation grounds our formalism in observable consequences. Saying an agent "believes climate change is serious" means they assign probability mass to action policies involving low-carbon choices, supporting climate policy, adjusting consumption patterns. The distribution $P_i(\theta)$ encodes the full behavioral profile implied by their belief state.

Messages similarly carry action policy distributions. A message m proposes—implicitly or explicitly—how agents should behave: $P_m(\theta)$ is the distribution over action policies consistent with accepting the message. For "care for future generations," P_m is diffuse and low-amplitude: many behavioral patterns compatible with this value, none demanded with extreme specificity. For "implement \$50/ton carbon tax," P_m is peaked and high-amplitude: accepting this message concentrates probability mass on specific political actions.

The amplitude of a distribution captures behavioral commitment. Mathematically, we can measure this through the concentration of the distribution:

$$\text{Amplitude}(P) = \int_\Theta P(\theta)^2 d\theta \tag{4}$$

High amplitude means probability mass is concentrated (peaked distribution, specific behavioral demands). Low amplitude means probability mass is spread (diffuse distribution, general reorientation).

When agent i receives message m, the cognitive dissonance of adoption is the information-theoretic distance between their current distribution and the message's proposed distribution:

$$\Delta_i(m) = D_{KL}(P_i \| P_m) = \int_\Theta P_i(\theta) \log \frac{P_i(\theta)}{P_m(\theta)} d\theta \tag{5}$$

This KL divergence quantifies precisely how much probability mass must be relocated in action policy space. Small $\Delta_i(m)$ means the message proposes action policies already assigned reasonable probability by agent i—adoption requires only minor adjustment. Large $\Delta_i(m)$ means the message proposes action policies currently assigned negligible probability—adoption requires major reorganization.

2.4 The Frequency-Distribution Connection

The profound insight linking spectral and distributional views is that frequency content in semantic space corresponds to distributional properties in action policy space. This connection emerges because smooth patterns on the semantic graph arise precisely when messages propose similar action distributions.

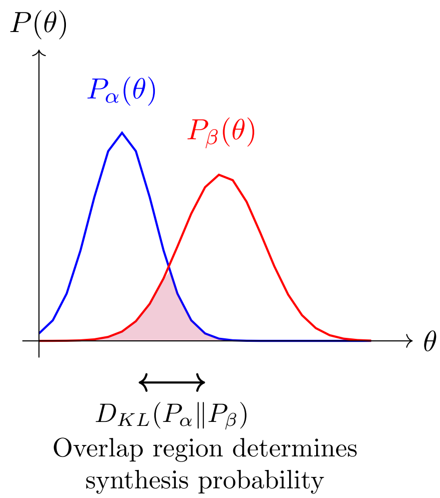

Figure 3: Message collision as distribution overlap. Messages m_α and m_β carry probability distributions P_α and P_β over belief space. The shaded overlap region indicates compatibility—large overlap (small KL divergence) enables synthesis, producing offspring message m_γ whose distribution integrates both parents. The KL divergence $D_{KL}(P_\alpha \| P_\beta)$ quantifies the information-theoretic distance that determines synthesis probability.

Proposition 1 (Frequency-Distribution Correspondence)**.** *A message with low-frequency content on the semantic graph G_M induces, on average, small KL divergence from agent beliefs. A message with high-frequency content induces large average KL divergence.*

Sketch. Low-frequency eigenvectors ϕ_k (small λ_k) vary smoothly: semantically similar messages have similar values. This means messages with primarily low-frequency content propose action distributions that change gradually across semantic neighborhoods. If agents' beliefs are also relatively smooth (reflecting that nearby positions in semantic space generally hold related action commitments), then low-frequency messages encounter priors that aren't radically different from what they propose, yielding small D_{KL}.

Conversely, high-frequency eigenvectors oscillate rapidly: nearby messages in semantic space can have opposite signs. Messages with high-frequency content thus propose action distributions that vary sharply even between semantically related positions. When such messages encounter agent beliefs, they typically propose distributions far from current commitments, yielding large D_{KL}. □

This correspondence explains our empirical observations. "Care for future generations" has low frequency because it proposes action distributions (gentle, diffuse priority shifts) that vary smoothly across semantic space—almost everyone's current beliefs include *some* future-regarding component, so the KL divergence remains small. "Specific carbon tax" has high frequency because it proposes distributions (concentrated on particular policy actions) that vary sharply—agents committed to market solutions versus government intervention have radically different priors, making the KL divergence large for most positions.

The beauty of this correspondence is that we can work in whichever representation proves more convenient. Frequency content (geometric, visualizable) and KL divergence (information-theoretic, quantitative) describe the same underlying phenomenon from different vantage points.

2.5 Message Propagation Dynamics

Messages spread through the agent graph with rates determined by their frequency content via the cognitive dissonance mechanism. Let $\pi_m(t)$ denote message m's prevalence—the fraction of agents who have adopted or been exposed to this message. Its evolution follows:

$$\frac{d\pi_m}{dt} = \underbrace{\sum_{i \in V} r_i(m) \cdot (1 - \pi_m)}_{\text{spread rate}} - \underbrace{\delta_m \cdot \pi_m}_{\text{decay rate}} \tag{6}$$

The adoption rate $r_i(m)$ captures how readily agent i accepts message m upon exposure. This rate decreases exponentially with cognitive dissonance:

$$r_i(m) = r_0 \cdot e^{-\beta \Delta_i(m)} = r_0 \cdot e^{-\beta D_{KL}(P_i \| P_m)} \tag{7}$$

The parameter β reflects agents' resistance to cognitive dissonance—how much they discount messages requiring large belief updates. The exponential form captures a fundamental psychological reality: acceptance probability drops dramatically as demanded updates grow.

This creates immediate frequency-dependent selection. By Proposition **??**, low-frequency messages induce small average Δ_i, yielding large $r_i(m)$ for most agents—they spread rapidly. High-frequency messages induce large average Δ_i, yielding small $r_i(m)$—they spread slowly, confined to communities whose beliefs already occupy nearby regions of action policy space.

The decay term captures forgetting, obsolescence, or displacement by competing messages:

$$\delta_m = \delta_0 + \sum_{m' \neq m} \alpha_{mm'} \pi_{m'} \tag{8}$$

Messages compete for limited cognitive resources. High-salience messages (those distinctive from background) displace similar but less interesting messages.

2.6 Salience and the Intermediate Frequency Sweet Spot

Spread rate alone does not determine message success. Agents must first *notice* messages before deciding whether to adopt them. Salience—the perceptual distinctiveness of a message—depends on how much it differs from the ambient discourse:

Definition 4 (Message Salience)**.** *The salience $S(m)$ of message m relative to current discourse is its expected KL divergence from prevalent messages:*

$$S(m) = \sum_{m' \in M} D_{KL}(P_m \| P_{m'}) \cdot \pi_{m'} \tag{9}$$

Messages with purely low-frequency content have low salience: they propose action distributions barely distinguishable from what agents already encounter. The KL divergences $D_{KL}(P_m \| P_{m'})$ remain small across prevalent messages m'. While such messages spread easily once noticed, they struggle to capture attention.

Messages with extreme high-frequency content also face salience challenges despite their sharp distinctiveness. They're so far from the space of active discourse that they lack entry points—agents can't connect them to ongoing conversations. The KL divergences are large but lack the structure that enables engagement.

Intermediate-frequency messages optimize this tradeoff. They're distinctive enough (moderate KL divergence from background) to capture attention while smooth enough to enable cognitive integration. They propose action distributions that are *interestingly different* from current commitments without being incomprehensibly alien.

We can formalize message fitness as combining spread and salience:

$$F(m) = \underbrace{e^{-\beta\bar{\Delta}(m)}}_{\text{spreadability}} \times \underbrace{S(m)}_{\text{salience}} - \delta_m \tag{10}$$

where $\bar{\Delta}(m) = \mathbb{E}_i[\Delta_i(m)]$ is the average cognitive dissonance across agents. This fitness function is non-monotonic in frequency: very low frequencies maximize spreadability but minimize salience, very high frequencies maximize salience but minimize spreadability. The optimum lies at intermediate frequencies where both factors contribute meaningfully.

3 Discourse Dynamics and Equilibria

3.1 The Frequency-Spread Tradeoff

Our fundamental prediction concerns how spread rate varies with frequency content:

Theorem 1 (Frequency-Spreadability Relationship)**.** *For agents with smoothly varying beliefs over the semantic graph, message spread rate $\bar{r}(m)$ decreases monotonically with the message's center frequency:*

$$\bar{k}_m = \frac{\sum_k k \cdot |c_{m,k}|^2}{\sum_k |c_{m,k}|^2} \tag{11}$$

where $\{c_{m,k}\}$ are the spectral coefficients of message m in the eigenbasis of G_M.

Sketch. Messages with high center frequency $\bar{k}_m$ have significant projection onto high-eigenvalue eigenvectors ϕ_k (large λ_k). These eigenvectors vary sharply across the semantic graph. When message m has high-frequency content, it proposes action distributions P_m that differ substantially from the distributions P_i of typical agents, because agents' beliefs—being relatively smooth over semantic space—don't follow the sharp oscillations of high-frequency eigenvectors.

By Proposition **??**, this yields large average KL divergence: $\bar{\Delta}(m) = \mathbb{E}_i[D_{KL}(P_i \| P_m)]$ grows with $\bar{k}_m$. By Equation (**??**), adoption rate decreases exponentially with KL divergence: $\bar{r}(m) = r_0 e^{-\beta\bar{\Delta}(m)}$. Composing these relationships yields the desired monotonic decrease. □

This theorem formalizes our earlier intuition. Low-frequency messages ("care for future generations," high projection on ϕ_0 and other small-λ eigenvectors) create smooth patterns requiring small belief updates, hence spreading rapidly. High-frequency messages ("specific policy with exact dates," high projection on large-λ eigenvectors) create sharp patterns requiring large belief updates, hence spreading slowly.

The spread rate doesn't merely correlate with frequency—it depends exponentially on it through the cognitive dissonance mechanism. A message with twice the center frequency doesn't spread half as fast; it spreads exponentially slower.

3.2 Equilibrium Discourse Structure

The interplay between frequency-dependent spread and salience creates predictable equilibrium structures. At equilibrium, message prevalence π_m stabilizes where spread balances decay: $\frac{d\pi_m}{dt} = 0$.

Proposition 2 (Intermediate Frequency Dominance)**.** *At equilibrium, the prevalence-weighted frequency distribution exhibits a peak at intermediate frequencies. Specifically, if we weight messages by their prevalence and engagement, the resulting distribution over eigenvalues λ_k takes the form:*

$$\rho(\lambda) \propto \lambda^{\alpha} e^{-\beta\lambda} \tag{12}$$

for parameters $\alpha, \beta > 0$ determined by the salience-spread tradeoff.

This gamma-like distribution captures the sweet spot phenomenon. The power law factor λ^{α} suppresses very low frequencies (poor salience), while the exponential decay $e^{-\beta\lambda}$ suppresses very high frequencies (poor spread). The peak—corresponding to maximum sustainable discourse engagement—occurs at:

$$\lambda^* = \frac{\alpha}{\beta} \tag{13}$$

This prediction is testable: by analyzing discourse over time and computing the spectral content of messages weighted by engagement metrics, we should observe this characteristic peak at intermediate eigenvalues rather than monotonic distributions.

3.3 Polarization as Spectral Fragmentation

Perhaps the most striking prediction concerns polarization—the phenomenon where communities lose ability to communicate effectively despite maintaining internal coherence. Spectral thinking reveals polarization as *loss of low-frequency overlap.*

Initially, diverse communities may disagree on high-frequency specifics (particular policies, interpretations of events) while maintaining low-frequency alignment (shared values, common ground on fundamentals). This low-frequency overlap, captured by agents' probability distributions over action policies having small mutual KL divergence in the low-frequency subspace, enables communication: messages can be constructed that, while perhaps intermediate or high frequency overall, remain low frequency relative to the shared substrate.

As communities drift apart—perhaps through selective exposure, filter bubbles, or deliberate wedge issues—they first lose high-frequency alignment (disagreement on specifics increases). But the system remains communicable as long as low-frequency overlap persists. The critical transition occurs when even the low-frequency structure diverges.

Theorem 2 (Polarization Phase Transition)**.** *There exists a critical threshold τ_c in the low-frequency spectral overlap:*

$$O_{AB}^{low} = \sum_{k:\lambda_k<\Lambda} \langle P_A, \phi_k\rangle \langle P_B, \phi_k\rangle \tag{14}$$

below which the system undergoes polarization. Here P_A, P_B are the aggregate action policy distributions of communities A and B, and Λ is a cutoff defining the low-frequency subspace.

When $O_{AB}^{low} < \tau_c$, even messages with primarily low-frequency content induce large mutual KL divergence, causing cross-community communication to collapse.

The mechanism is insidious. As low-frequency overlap erodes, messages that previously bridged communities (intermediate frequency, moderate KL divergence) become effectively high frequency across the divide. An idea that would have required moderate belief updates from both sides now requires large updates, triggering exponential resistance. Communities enter echo chambers not because they explicitly choose to, but because the cognitive dissonance of cross-community messages exceeds adoption thresholds.

This explains empirical patterns in political polarization: partisans report increasing difficulty understanding opponents' positions, not just disagreeing with them. The loss isn't merely in high-frequency policy specifics but in the low-frequency conceptual foundations—what counts as evidence, which sources to trust, which values to prioritize. When these diverge sufficiently, genuine communication becomes information-theoretically expensive.

3.4 Multi-Timescale Structure

The dynamics exhibit natural timescale separation that simplifies analysis. Message adoption and spread occur on fast timescales (hours to days in online contexts), modulated by the exponential factor $e^{-\beta\Delta_i(m)}$ which creates rapid accept/reject decisions. Agent belief distributions evolve on slow timescales (weeks to months), accumulating evidence from multiple message exposures and personal experiences. The semantic graph structure G_M is intermediate, evolving as new topics emerge but remaining relatively stable compared to message flows.

This separation enables a hierarchical approximation:

1. **Fast dynamics**: Message propagation reaches quasi-equilibrium given current agent beliefs and semantic structure
2. **Intermediate dynamics**: Semantic graph topology adjusts as discourse focus shifts between topics
3. **Slow dynamics**: Agent belief distributions evolve through accumulated message exposure

At the slow timescale, we can treat message prevalence as approximately equilibrated, with distributions π_m determined by the frequency-dependent fitness function. This equilibrium prevalence then drives gradual shifts in agent beliefs, which in turn modify the effective frequency content of messages (as what counts as "smooth" vs "sharp" depends on the belief landscape).

The multi-timescale structure also explains why rapid message churn in public discourse can mask stable underlying patterns. At the fast timescale, individual messages appear and disappear chaotically. But when we measure the *spectral content* of discourse—the frequency distribution weighted by engagement—we find stable structure reflecting the fundamental dynamics of cognitive dissonance and salience.

4 Empirical Validation Protocols

The framework makes several concrete predictions amenable to rigorous testing using modern computational methods and large-scale discourse data. We detail specific protocols for the three core predictions.

4.1 Protocol 1: Frequency Predicts Spread Rate

Hypothesis: Message spread rate decreases with spectral center frequency, controlling for network effects, source characteristics, and temporal confounds.

Data Requirements:

- Social media cascade data (Twitter retweets, Reddit cross-posts) with full propagation networks
- Scientific citation cascades from arXiv, Semantic Scholar, or similar
- Minimum 10,000 messages with complete diffusion histories

Experimental Protocol:

1. **Semantic graph construction**: Build message graph G_M where edges connect semantically similar messages using:
   - Sentence-BERT embeddings to compute pairwise semantic similarity
   - Threshold at 90th percentile to create sparse graph
   - Verify graph connectivity (exclude disconnected components)
2. **Spectral decomposition**: Compute normalized graph Laplacian $\mathcal{L} = D^{-1/2} L_M D^{-1/2}$ and extract eigenvectors $\{\phi_k\}$ and eigenvalues $\{\lambda_k\}$
3. **Frequency content assignment**: For each message m:
   - Project message embedding onto eigenbasis: $c_k = \langle m, \phi_k \rangle$
   - Compute center frequency: $\bar{k}_m = \sum_k k|c_k|^2 / \sum_k |c_k|^2$
   - Normalize to [0,1] scale for interpretability
4. **Spread rate measurement**: For each message, compute:
   - Peak velocity: $v_m = \max_t(d\pi_m/dt)$
   - Time to half-saturation: $\tau_{1/2}$ when π_m reaches 50% of eventual value
   - Total reach: final prevalence $\pi_m(\infty)$ normalized by network position
5. **Control variables**:
   - Source authority: follower count, PageRank centrality
   - Network position: betweenness, closeness centrality
   - Content features: length, media type, linguistic complexity
   - Temporal effects: time of day, day of week, time-fixed effects
6. **Statistical analysis**: Hierarchical regression with random effects for users:

$$\log(v_m) = \beta_0 + \beta_1 \bar{k}_m + \beta_2 S(m) + \mathbf{Z}'\boldsymbol{\gamma} + \epsilon \tag{15}$$

   where $\mathbf{Z}$ includes controls. Test $H_0 : \beta_1 = 0$ vs $H_a : \beta_1 < 0$
7. **Robustness checks**:
   - Within-user analysis: compare messages from same sources
   - Propensity score matching: balance treatment and control on observables
   - Instrumental variable: use exogenous variation in semantic graph structure

Expected Results:

- Coefficient $\beta_1 < 0$ significant at $p < 0.01$
- Effect size: 0.1 increase in normalized center frequency predicts 15-25% decrease in peak velocity
- Relationship robust to control inclusion, strongest for messages crossing community boundaries

Alternative Explanations to Eliminate:

- *Audience size confound*: High-frequency messages might have smaller potential audience. Control: normalize reach by potential audience using network distance metrics
- *Source selection*: Influential sources might preferentially share low-frequency content. Control: within-source comparison, source fixed effects
- *Reverse causality*: Slow spread might cause higher measured frequency if late adopters differ. Control: use only first 24 hours of cascade for frequency measurement

4.2 Protocol 2: Intermediate Frequencies Dominate Engagement

Hypothesis: In sustained discourse, intermediate-frequency messages account for disproportionate engagement compared to their prevalence.

Data Requirements:

- Long-running online communities (Reddit, specialized forums) with 6+ months of continuous activity
- Full comment threads, voting data, user engagement metrics
- Minimum 50,000 messages across diverse topics

Experimental Protocol:

1. **Community identification**: Use community detection (Louvain modularity) to identify stable discussion communities within platform
2. **Temporal windowing**: Divide data into 2-week windows, compute separate message graphs for each window to capture evolving discourse
3. **Engagement weighting**: For each message m, measure total engagement:
$$E(m) = w_1 \cdot \text{replies} + w_2 \cdot \text{upvotes} + w_3 \cdot \text{dwell time} \tag{16}$$
with weights w_i estimated from user surveys about engagement intensity
4. **Distribution computation**: Calculate two frequency distributions:
    - Prevalence distribution: $\rho_{\text{prev}}(\lambda)$ = fraction of messages with center frequency λ
    - Engagement distribution: $\rho_{\text{engage}}(\lambda)$ = engagement-weighted fraction
5. **Peak identification**: Fit gamma distribution (Proposition **??**) to engagement distribution:
$$\rho(\lambda) = A\lambda^\alpha e^{-\beta\lambda} \tag{17}$$
using maximum likelihood estimation
6. **Comparison tests**:
    - Kolmogorov-Smirnov test: is ρ_{engage} different from ρ_{prev}?
    - Peak comparison: is peak of ρ_{engage} at higher λ than peak of ρ_{prev}?
    - Concentration measure: what fraction of engagement goes to top 20% intermediate-frequency messages?
7. **Longitudinal tracking**: Monitor how frequency distributions evolve over discourse lifetime:

- Do communities converge toward predicted equilibrium distribution?
- What is the relaxation time: how quickly does convergence occur?
- Do external shocks (major events) perturb the distribution, and how does it recover?

Expected Results:

- Engagement distribution peaks at $\lambda^* \approx 0.35$-0.50 of maximum eigenvalue (intermediate range)
- Prevalence distribution either monotonically decreasing or peaked at lower frequencies
- Gamma fit superior to alternatives (power law, lognormal): $\Delta\text{AIC} > 10$
- Top 20% of intermediate-frequency messages capture 60-70% of total engagement

4.3 Protocol 3: Polarization Through Low-Frequency Divergence

Hypothesis: Community polarization corresponds to declining low-frequency spectral overlap, with phase transition when overlap drops below critical threshold.

Data Requirements:

- Political discourse during election cycles (18+ months of Twitter, Reddit data)
- Communities undergoing schism (subreddit splits, forum migrations)
- Control: stable communities maintaining cross-group interaction

Experimental Protocol:

1. **Community pair selection**: Identify pairs of communities that:
   - Initially show substantial cross-community interaction (shared users, cross-posts)
   - Subsequently polarize (declining interaction, increasing mutual hostility)
   - Maintain internal coherence (stable within-community engagement)
2. **Aggregate belief distributions**: For each community C and time window t:
   - Aggregate all messages into community corpus
   - Compute semantic embedding of corpus: $e_C^{(t)}$
   - Treat embedding as proxy for aggregate action policy distribution
   - Project onto message graph eigenbasis: $P_C^{(t)} = \sum_k c_{C,k}^{(t)} \phi_k$
3. **Spectral overlap trajectory**: Compute low-frequency overlap over time:
$$O^{(t)}(A, B) = \frac{\sum_{k:\lambda_k < \Lambda} c_{A,k}^{(t)} c_{B,k}^{(t)}}{\sqrt{\sum_{k:\lambda_k < \Lambda} (c_{A,k}^{(t)})^2} \sqrt{\sum_{k:\lambda_k < \Lambda} (c_{B,k}^{(t)})^2}} \tag{18}$$
   where Λ is the 10th percentile eigenvalue (defining low-frequency subspace)
4. **KL divergence evolution**: Compute information-theoretic distance:
$$D_{KL}^{(t)}(A, B) = D_{KL}(P_A^{(t)} \| P_B^{(t)}) \tag{19}$$
   using aggregate distributions

5. **Interaction rate measurement**: Track cross-community interaction intensity:

$$I^{(t)}(A, B) = \frac{\text{cross-community posts}^{(t)}}{\text{total posts}^{(t)}} \tag{20}$$

6. **Phase transition identification**: Test for critical threshold:
   - Hypothesis: $I^{(t)}$ drops sharply when $O^{(t)} < \tau_c$
   - Use changepoint detection to identify critical overlap value
   - Test whether drop is sharp (phase transition) vs gradual (continuous evolution)

7. **Control comparison**: For each polarizing pair, match with stable control pair:
   - Similar initial overlap and interaction levels
   - Similar topic domains and community sizes
   - Difference-in-differences: $\Delta O^{(t)}_{\text{polarized}} - \Delta O^{(t)}_{\text{control}}$

8. **Causal inference**: Use instrumental variables or natural experiments:
   - Platform design changes affecting cross-community exposure
   - Major external events forcing interaction
   - Policy interventions promoting bridging

Expected Results:

- Low-frequency overlap $O^{(t)}$ declines 40-60% during polarization period
- High-frequency components show earlier divergence but don't predict interaction collapse
- Critical threshold: when $O^{(t)} < 0.30$, cross-community interaction drops by 5-10x within 2-4 weeks
- Phase transition observable: sharp drop in $I^{(t)}$ rather than gradual decline
- Control communities maintain $O^{(t)} > 0.50$ throughout observation period

4.4 Implementation Resources

To facilitate replication and extension, we will provide:

- **Open-source Python library**: Implements spectral message analysis, including graph Laplacian computation, frequency decomposition, and statistical testing
- **Pre-processed datasets**: Validated message graphs from Twitter, Reddit, arXiv with ground-truth propagation data
- **Interactive visualizations**: Tools for exploring spectral content and discourse dynamics
- **Standardized protocols**: Detailed code templates for each validation study
- **Replication materials**: Complete analysis scripts and documentation

These resources will enable other researchers to validate our predictions, extend the framework to new domains, and develop improved methods for spectral discourse analysis.

5 Discussion and Future Directions

5.1 Theoretical Connections

The spectral framework bridges several established research traditions, revealing them as different perspectives on unified phenomena.

Active Inference and Free Energy: The cognitive dissonance measure $\Delta_i(m) = D_{KL}(P_i \| P_m)$ is precisely the variational free energy cost of updating beliefs in active inference (**??**). Our framework extends active inference from individual agents to cultural dynamics: communities collectively minimize free energy by coordinating on low-frequency semantic structure (broad value alignment) while tolerating high-frequency variation (specific policy disagreements). This suggests viewing discourse as distributed Bayesian inference where messages serve as observations that agents integrate to update collective beliefs.

Graph Signal Processing: Spectral decomposition of semantic graphs applies established graph signal processing techniques (**?**) to cultural content. Our contribution lies in connecting spectral properties to memetic fitness through the action policy interpretation: frequency content determines not merely abstract graph-theoretic properties but concrete behavioral implications. This grounding in action transforms graph signal processing from descriptive mathematics into predictive theory.

Complex Contagion: The exponential suppression of high-frequency message adoption generalizes complex contagion models (**?**). Where complex contagion treats adoption thresholds as exogenous parameters, we derive them from information-theoretic principles: thresholds emerge naturally from cognitive dissonance, varying with message frequency content and agent belief distributions. This provides microfoundations for threshold phenomena in social spreading.

Cultural Evolution: Our framework formalizes cultural selection pressures through frequency-dependent fitness (**?**). Messages undergo selection not merely for intrinsic appeal but for their position in frequency space relative to population belief distributions. This explains phenomena like belief perseverance and ideological path-dependence: communities develop characteristic frequency profiles that create selection pressures favoring compatible messages.

5.2 Open Questions

Several important questions warrant future investigation:

Strategic Message Design: How do sophisticated actors exploit spectral structure to craft maximally effective messages? Understanding adversarial dynamics—politicians optimizing rhetoric for target demographics, propagandists maximizing spread while minimizing fact-checking resistance—could reveal new approaches to detecting and countering manipulation. The framework predicts that effective propaganda operates at intermediate frequencies: specific enough to coordinate action, smooth enough to spread broadly.

Network Topology Effects: How does agent graph structure G_A modulate message dynamics? Small-world topologies with bridging ties might facilitate cross-frequency collision by exposing agents to diverse message types. Clustered networks with strong community structure might amplify polarization by creating echo chambers where local frequency distributions diverge. The interaction between network structure and spectral dynamics deserves systematic analysis.

Multi-Modal Communication: Real messages combine linguistic, visual, and emotional content. Different modalities might carry different frequency signatures: images often convey low-frequency emotional appeals while text carries higher-frequency specific claims. Understanding how multi-modal content projects onto semantic space could explain why certain message formats (memes, videos, infographics) spread effectively despite cognitive complexity.

Temporal Dynamics: We've focused on quasi-equilibrium dynamics, but transient phenomena matter. How do discourse communities respond to sudden shocks (major news events, platform policy changes)? Do perturbations induce rapid frequency shifts or merely temporary prevalence changes that quickly relax back to equilibrium? Can we predict discourse evolution using dynamical systems techniques?

Institutional Design: What organizational structures optimize collective intelligence through spectral principles? Deliberation platforms might implement frequency-aware algorithms: surface intermediate-frequency messages for maximum engagement, detect low-frequency divergence as early warning for polarization, orchestrate controlled collisions between compatible frequency bands. Translating spectral theory into practical design remains an open challenge.

5.3 Implications for Collective Intelligence

The spectral framework suggests reconceptualizing collective intelligence design around *spectral health*—maintaining appropriate frequency distributions rather than optimizing single metrics.

Healthy discourse requires strong low-frequency alignment (shared values, common conceptual foundations) enabling coordination, combined with rich high-frequency diversity (varied specific perspectives, creative disagreement) enabling innovation. Pathology arises from two failure modes: loss of low-frequency overlap (polarization, making communication information-theoretically expensive) or suppression of high-frequency variation (groupthink, eliminating novel ideas).

This suggests design principles for information ecosystems:

1. **Preserve common ground**: Make low-frequency agreements visible even amid high-frequency disagreement. Platforms could highlight areas of consensus, surface shared values, create regular rituals of collective affirmation that maintain spectral overlap.
2. **Optimize bridging frequency**: Focus deliberation on intermediate-frequency topics where productive synthesis is possible. Rather than seeking lowest-disagreement topics (too low frequency, unsalient) or most controversial (too high frequency, unresolvable), identify structured questions that matter yet remain discussable.
3. **Manage frequency transitions**: Help agents navigate from high-frequency observations (specific experiences, novel data) toward intermediate-frequency synthesis (structured but spreadable insights). This might involve facilitated workshops that progressively smooth rough observations into resonant principles.
4. **Monitor spectral divergence**: Track low-frequency overlap between communities as early warning indicator for polarization. Intervene before crossing critical thresholds by creating structured cross-community interactions focused on rediscovering shared foundations.
5. **Design for frequency diversity**: Avoid algorithms that exclusively amplify either low-frequency (bland consensus) or high-frequency (outrage and controversy) content. Healthy discourse requires the full spectrum, with appropriate prevalence at each frequency band.

These principles don't solve collective intelligence challenges—they provide a framework for thinking about solutions. The spectral perspective reveals that many apparent trade-offs (consensus vs diversity, coordination vs innovation, agreement vs truth-seeking) reflect inappropriate frequency optimization. By understanding discourse as possessing spectral structure, we can design interventions that work with, rather than against, the natural dynamics of idea propagation.

6 Conclusion

We have introduced a mathematical framework for understanding message propagation through spectral structure—the frequency content that determines how ideas resonate with and transform agent beliefs. By viewing messages simultaneously through geometric (patterns on semantic graphs), information-theoretic (probability distributions over action policies), and behavioral (demanded updates to behavioral patterns) lenses, we reveal these as unified perspectives on a single phenomenon.

The framework makes testable predictions: spread rates decrease with frequency content, equilibrium discourse exhibits intermediate-frequency dominance, polarization corresponds to low-frequency divergence. These predictions are amenable to rigorous validation using modern computational methods and large-scale discourse data. The protocols we've detailed provide concrete paths toward empirical confirmation or refutation.

Beyond specific predictions, spectral thinking offers a new way of approaching collective intelligence design. Rather than treating discourse as chaotic information flow requiring top-down management, we can understand it as possessing natural frequency structure—patterns that emerge from the interaction between cognitive dissonance and semantic topology. By working with this structure rather than against it, we might design information ecosystems that enhance rather than degrade our collective capacity for understanding.

The journey from acoustic metaphors ("ideas resonate like sound waves") to rigorous mathematics (graph Laplacians, KL divergence, spectral decomposition) to empirical predictions (frequency-spread relationships, polarization thresholds, bridging algorithms) demonstrates how deep analogies can ground formal theory. Frequency isn't merely a useful metaphor—it's the natural language for describing how patterns propagate through networked populations of belief-updating agents.

As artificial intelligence systems increasingly participate in human discourse, bringing their own characteristic frequency profiles and belief-updating dynamics, understanding spectral structure becomes not merely intellectually interesting but practically essential. The collision chamber of culture is entering a new phase where multiple forms of intelligence interact, and we need mathematics adequate to this complexity—mathematics that respects information-theoretic constraints, captures multi-scale dynamics, and connects individual cognition to collective outcomes. This framework offers one path toward that understanding.

Acknowledgments

We thank Peter Thestrup Waade for critical feedback that substantially improved the exposition, particularly the emphasis on grounding frequency content in action policy updates. [Additional acknowledgments to be added.]

References

Dawkins, R. (1976). *The Selfish Gene*. Oxford University Press.

Blackmore, S. (1999). *The Meme Machine*. Oxford University Press.

Centola, D. (2010). The spread of behavior in an online social network experiment. *Science*, 329(5996), 1194-1197.

Friston, K. (2009). The free-energy principle: a rough guide to the brain? *Trends in Cognitive Sciences*, 13(7), 293-301.

Ramstead, M. J., Badcock, P. B., & Friston, K. J. (2018). Answering Schrödinger's question: A free-energy formulation. *Physics of Life Reviews*, 24, 1-16.

Shuman, D. I., Narang, S. K., Frossard, P., Ortega, A., & Vandergheynst, P. (2013). The emerging field of signal processing on graphs. *IEEE Signal Processing Magazine*, 30(3), 83-98.

Mesoudi, A. (2011). *Cultural Evolution: How Darwinian Theory Can Explain Human Culture and Synthesize the Social Sciences.* University of Chicago Press.