Spectral Byzantine Consensus: Anisotropic Aggregation on Continuous Semantic Fields

Cyfrpunk Raelizde
Peak&Tower Research
research@peakandtower.io

November 17, 2025

Abstract

We present a mathematical framework for achieving Byzantine fault-tolerant consensus on continuous functions defined over graph-structured domains. Traditional Byzantine consensus protocols operate on discrete state (transaction ordering, block validation), while distributed AI systems require consensus on *continuous semantic representations*: neural network embeddings, probability distributions, or geometric fields with inherent manifold structure.

Our protocol operates in the *spectral domain* of the graph Laplacian, representing distributed state as coefficients in an eigenbasis. We combine three key components: (1) *mode-wise geometric median aggregation* with provable Byzantine robustness bounds from robust statistics, (2) *anisotropic heat flow dynamics* that preferentially preserve low-frequency (global) structure while damping high-frequency (local) noise, and (3) *spectral coherence certificates* that detect geometric inconsistencies under Dirichlet energy constraints.

We prove that for Byzantine weight fraction $\rho_w < 1/3$, the protocol achieves consensus with per-mode error bounded by a constant times $\sigma_k \sqrt{\rho_w/(1-\rho_w)}$, where σ_k is the honest noise variance in mode k. The convergence rate is governed by a weighted Lyapunov function with contraction coefficient $\gamma \approx \alpha_0 \lambda_\star/2$, where $\lambda_\star$ is the spectral gap of the honest communication graph.

Our framework bridges three research communities: (1) Byzantine fault tolerance (extending discrete consensus to continuous fields), (2) graph signal processing (adding adversarial robustness to spectral methods), and (3) distributed machine learning (enabling federated learning on geometric data with provable Byzantine resilience). We provide fast approximation algorithms using Chebyshev polynomials, avoiding explicit eigendecomposition while maintaining theoretical guarantees.

1 Introduction

1.1 Motivation: From Discrete to Continuous Consensus

Classical Byzantine fault-tolerant (BFT) consensus protocols [4, 3, 10] solve a fundamental problem: achieving agreement among distributed agents when some fraction may behave arbitrarily. These protocols excel at ordering discrete events—transactions in a blockchain, state machine transitions in replicated services.

However, emerging distributed AI applications operate on *continuous semantic representations*:

- **Multi-agent world modeling**: Autonomous vehicles, robots, or drones build shared representations of 3D environments, represented as neural radiance fields, occupancy grids, or semantic embeddings.
- **Federated learning on geometric data**: Distributed training of neural networks on graph-structured data (molecules, social networks, knowledge graphs) requires aggregating gradients or model parameters that live in high-dimensional spaces with geometric structure.
- **Decentralized AI safety**: Validators in a blockchain-based AI inference network must reach consensus on model outputs (probability distributions, embeddings) without trusting any single node.

These applications motivate consensus protocols that:

1. Handle *continuous state spaces* (vectors in $\mathbb{R}^d$, functions on manifolds) rather than discrete choices.
2. Respect *geometric structure* (manifold constraints, smoothness priors, locality).
3. Provide *Byzantine robustness* with mathematical guarantees.
4. Scale to *high-dimensional representations* (thousands to millions of parameters).

1.2 Our Approach: Spectral Byzantine Consensus

We propose representing distributed state in the *spectral domain* of a learned graph Laplacian. Consider a set of agents $\mathcal{A}$ maintaining beliefs $f^{(a)} : V \to \mathbb{R}$ over a graph $G = (V, E)$ (e.g., a discretized spatial domain, a knowledge graph, or a neural network's feature space). Instead of exchanging raw function values, agents communicate *spectral coefficients*:

$$c_k^{(a)} = \phi_k^\top f^{(a)} = \sum_{i \in V} \phi_k(i) f^{(a)}(i) \in \mathbb{R}$$

where $\{\phi_k\}$ are the eigenvectors of the normalized graph Laplacian L.

Key insight: The eigenbasis $\{\phi_k\}$, ordered by eigenvalue λ_k, forms a *frequency decomposition*:

- Low modes ($\lambda_k \ll 1$): Global, smooth structure (constant or slowly varying over the graph).
- High modes ($\lambda_k \approx 2$): Local, rapid oscillations (high-frequency detail).

This spectral representation enables three critical capabilities:

1. Anisotropic Byzantine Resistance By applying geometric median aggregation *mode-by-mode*, we can prove that Byzantine influence is mode-dependent. Crucially, low-frequency modes (which capture global semantic structure) are *harder for adversaries to corrupt* because honest agents' coefficients naturally cluster. This allows us to prioritize preserving consensus on global structure over local detail.

2. Fast Convergence via Heat Flow We design update rules that mimic discrete heat diffusion:

$$c_{i,k}^{(t+1)} = (1 - \alpha_k)c_{i,k}^{(t)} + \alpha_k \tilde{c}_{i,k}^{(t)}$$

where $\alpha_k = \eta\lambda_k$ couples the damping rate to eigenvalue. This creates *anisotropic convergence*: high-frequency modes (large λ_k) damp exponentially fast, while low-frequency modes (small λ_k) evolve slowly but robustly.

3. Geometric Coherence Certificates In the truncated K-mode spectral space, we can detect *geometric inconsistencies*: if adversarial agents claim two events e_1, e_2 are both valid but their spectral representations violate triangle inequality beyond what the energy budget allows, we have a computable proof of conflict. The effectiveness of this mechanism depends critically on external energy budget enforcement.

1.3 Contributions

Theoretical contributions.:

1. **Mode-wise Byzantine robustness** (Theorems 4.1, 4.3): We prove that weighted geometric median in scalar spectral space achieves error $\sigma_k(1 + 4\sqrt{\rho_w/(1-\rho_w)})$ for Byzantine weight fraction $\rho_w < 1/3$, extending Minsker's robust statistics to weighted settings. Trimmed mean aggregation achieves $\sigma_k(2 + 4\sqrt{\rho_w/(1-\rho_w)})$ with computational advantages.

2. **Anisotropic convergence analysis** (Theorems 5.4, 5.6): We establish exponential convergence to a Byzantine-perturbed equilibrium with mode-dependent rates $e^{-\eta\lambda_k t}$, and prove global contraction of a weighted Lyapunov function $V^{(t)} = \sum_k w_k V_k^{(t)}$ with $w_k = (1 + \lambda_k)^{-\beta}$. Convergence time is $O((\alpha_0\lambda_\star)^{-1}\log(1/\epsilon))$.

3. **Spectral fork detection** (Definition 6.4, Theorem 6.5): We derive computable certificates for geometric inconsistency based on Dirichlet energy constraints and triangle inequality violations in truncated spectral space. This mechanism requires external energy budget enforcement but provides a geometric foundation for fork detection distinct from voting-based approaches.

4. **Dynamic topology resilience** (Theorem 6.2): We extend all results to time-varying honest communication graphs under a uniform spectral gap assumption $\lambda_1(L_H^{(t)}) \geq \lambda_\star$.

Algorithmic efficiency.:

- **Chebyshev polynomial approximation**: We avoid explicit eigendecomposition (cost $O(n^3)$ for n-node graphs) by approximating spectral projections with Chebyshev polynomials of the Laplacian, achieving $O(K|E|)$ complexity for K modes and $|E|$ edges.
- **Fast geometric median**: We use Weiszfeld's algorithm with warm-starting across modes, achieving $O(m \log(1/\epsilon))$ iterations for m agents and tolerance ϵ.

Remark 1.1. While geometric medians, Chebyshev approximations, and spectral graph theory are classical tools, their combination into a Byzantine-robust consensus protocol operating natively in spectral space appears novel. Prior work applies spectral methods for post-hoc analysis (Section 8) or treats all dimensions uniformly in robust aggregation.

2 Preliminaries

2.1 Graph Laplacian and Spectral Decomposition

Definition 2.1 (Normalized Graph Laplacian)**.** Let $G = (V, E, w)$ be a weighted undirected graph with $n = |V|$ vertices and edge weights $w_{ij} \geq 0$. The degree matrix is $D = \text{diag}(d_1, \ldots, d_n)$ where $d_i = \sum_j w_{ij}$. The *normalized graph Laplacian* is:

$$L = I - D^{-1/2} W D^{-1/2}$$

where $W = (w_{ij})$ is the adjacency matrix.

The Laplacian is symmetric positive semidefinite with spectral decomposition:

$$L = \sum_{k=0}^{n-1} \lambda_k \phi_k \phi_k^\top$$

where:

- Eigenvalues: $0 = \lambda_0 < \lambda_1 \leq \cdots \leq \lambda_{n-1} \leq 2$
- Eigenvectors: $\{\phi_k\}_{k=0}^{n-1} \subset \mathbb{R}^n$ (orthonormal basis)

Definition 2.2 (Spectral Coefficients)**.** For a function $f : V \to \mathbb{R}$ represented as a vector $f \in \mathbb{R}^n$, the *spectral coefficient* for mode k is:

$$c_k := \phi_k^\top f = \sum_{i=1}^{n} \phi_k(i) f(i) \in \mathbb{R}$$

The function can be reconstructed as $f = \sum_{k=0}^{n-1} c_k \phi_k$.

Definition 2.3 (Dirichlet Energy)**.** The *Dirichlet energy* (or graph Laplacian quadratic form) of f is:

$$\mathcal{E}(f) := f^\top L f = \frac{1}{2} \sum_{(i,j)\in E} w_{ij} \left(\frac{f(i)}{\sqrt{d_i}} - \frac{f(j)}{\sqrt{d_j}} \right)^2 .$$

In spectral space: $\mathcal{E}(f) = \sum_{k=0}^{n-1} \lambda_k c_k^2$.

Intuition: The Dirichlet energy measures how much f varies across edges. Smooth (low-energy) functions have most energy in low-λ_k modes.

2.2 Chebyshev Polynomial Approximation

Computing eigenvectors $\{\phi_k\}$ directly via eigendecomposition costs $O(n^3)$, prohibitive for large graphs. We use Chebyshev polynomials to approximate spectral projections.

Definition 2.4 (Chebyshev Polynomials)**.** The Chebyshev polynomials of the first kind $\{T_k\}_{k=0}^{\infty}$ on $[-1, 1]$ satisfy:

$$T_0(x) = 1, \quad T_1(x) = x, \quad T_{k+1}(x) = 2xT_k(x) - T_{k-1}(x).$$

Lemma 2.5 (Chebyshev Approximation)**.** *Let L have eigenvalues in $[0, \lambda_{\max}]$. Define the scaled operator:*

$$\widetilde{L} = \frac{2}{\lambda_{\max}} L - I$$

with eigenvalues in $[-1, 1]$. For any $f \in \mathbb{R}^n$, the Chebyshev features:

$$u_k = T_k(\widetilde{L}) f$$

can be computed via the recurrence:

$$u_0 = f, \quad u_1 = \widetilde{L} f, \quad u_{k+1} = 2\widetilde{L} u_k - u_{k-1}$$

in $O(K|E|)$ time for K polynomials.

Moreover, $\left\| T_k(\widetilde{L}) \right\|_{\text{op}} \leq 1$, so noise in f does not amplify.

Remark 2.6. For the theoretical analysis in this paper we assume direct access to the spectral coefficients $c_k = \phi_k^\top f$, which can be obtained by standard Krylov methods in $O(K|E|)$ time. In practice, they can be approximated using the Chebyshev recurrence in Lemma 2.5; the approximation error does not affect the Byzantine robustness bounds, only the effective noise levels σ_k.

2.3 Byzantine-Robust Statistics

Definition 2.7 (Weighted Geometric Median)**.** For points $\{x_a\}_{a\in\mathcal{A}} \subset \mathbb{R}^d$ with weights $w_a > 0$ satisfying $\sum_a w_a = 1$, the *weighted geometric median* is:

$$m = \arg\min_{y\in\mathbb{R}^d} \sum_{a\in\mathcal{A}} w_a \left\| y - x_a \right\|_2 .$$

The geometric median is the multivariate generalization of the median with strong Byzantine robustness properties. We use the following result:

Theorem 2.8 (Minsker 2015, adapted)**.** *Let $\mathcal{A} = H \cup B$ where:*

- *Honest agents $a \in H$: $\|x_a - z\|_2 \leq \sigma$ for some ground truth $z \in \mathbb{R}^d$,*
- *Byzantine agents $b \in B$: arbitrary values,*
- *Byzantine weight fraction: $\rho_w := \sum_{b \in B} w_b < 1/3$,*

then the weighted geometric median m satisfies:

$$\|m - z\|_2 \leq C_{\text{med}} \sigma \sqrt{\frac{\rho_w}{1 - \rho_w}}$$

where $C_{\text{med}} = 4$ is a universal constant.

We apply this theorem *coordinate-wise* in spectral space (each c_k is a scalar in $\mathbb{R}$).

3 Protocol Description

3.1 System Model

Agents: A set $\mathcal{A} = H \cup B$ of n agents, where:

- Honest agents H: follow the protocol, represent ground truth with bounded noise;
- Byzantine agents B: controlled by an adversary, send arbitrary values.

Stake weights: Each agent a has weight $w_a > 0$ with $\sum_{a \in \mathcal{A}} w_a = 1$. We assume:

$$\rho_w := \sum_{a \in B} w_a < \frac{1}{3}$$

(Byzantine agents control less than 1/3 of total weight.)

Communication model: Agents communicate over an *honest subgraph* $G_H = (H, E_H)$ with Laplacian L_H. We assume G_H is connected (spectral gap $\lambda_1(L_H) > 0$). Byzantine agents can inject values into the aggregation but do not appear in the neighbor sampling distribution.

Semantic representation: Each agent a maintains a belief $f^{(a)} : V \to \mathbb{R}$ over a fixed domain graph $G = (V, E)$ (the "semantic space"). For honest agents:

$$\left\| f^{(a)} - f^\star \right\|_2 \leq \sigma_0$$

where $f^\star$ is the ground truth function.

3.2 Protocol State

At round t, each agent $i \in \mathcal{A}$ maintains:

$$c_i^{(t)} = (c_{i,0}^{(t)}, c_{i,1}^{(t)}, \ldots, c_{i,K-1}^{(t)}) \in \mathbb{R}^K$$

where $c_{i,k}^{(t)} \in \mathbb{R}$ is the coefficient for mode k.

Initialization: Each agent computes initial coefficients:

$$c_{i,k}^{(0)} = \phi_k^\top f^{(i)}(0)$$

using Chebyshev approximation (Lemma 2.5).

3.3 Protocol Round

At each round t, agent i executes:

Algorithm 1 Spectral Byzantine Consensus (Agent i, Round t)

for each mode $k \in \{0, \ldots, K-1\}$ **do**

  Gossip: Sample neighbors $N_i \subset H$ from lazy random walk on G_H.

  Collect: Gather coefficients $\{c_{j,k}^{(t)}\}_{j \in N_i}$ from sampled neighbors.

  Aggregate:

    Compute geometric median:

$$m_{i,k}^{(t)} = \arg\min_{y \in \mathbb{R}} \sum_{j \in N_i} w_j \left| y - c_{j,k}^{(t)} \right|.$$

    Sort neighbors by distance: $d_j = \left| c_{j,k}^{(t)} - m_{i,k}^{(t)} \right|$.

    Trim: Remove neighbors with largest d_j in descending order until cumulative removed weight $\geq \alpha_k = \eta \lambda_k$.

    Compute trimmed mean:

$$\tilde{c}_{i,k}^{(t)} = \sum_{j \in S_i} \widetilde{w}_j c_{j,k}^{(t)}$$

    with renormalized weights $\widetilde{w}_j$.

  Update:

$$c_{i,k}^{(t+1)} = (1 - \alpha_k) c_{i,k}^{(t)} + \alpha_k \tilde{c}_{i,k}^{(t)}.$$

end for

Parameters:

- K: number of spectral modes (typically $K \ll n$, e.g., $K = O(\log n)$);
- $\eta \in (0, 1/\lambda_{\max})$: heat flow damping rate;
- $\alpha_k = \eta \lambda_k$: mode-dependent step size;
- mode weights for Lyapunov analysis: $w_k = (1 + \lambda_k)^{-\beta}$ with $\beta > 0$.

4 Main Results I: Robust Aggregation

We first establish Byzantine robustness guarantees for the geometric median and trimmed mean aggregators applied to scalar spectral coefficients.

4.1 Weighted Geometric Median Bound

Theorem 4.1 (Scalar Geometric Median)**.** *Let $\{c_a\}_{a\in\mathcal{A}} \subset \mathbb{R}$ be scalar values with:*

- *honest agents $a \in H$: $|c_a - z| \le \sigma$;*
- *Byzantine agents $b \in B$: arbitrary values;*
- *stake weights: $w_a > 0$, $\sum_a w_a = 1$, and $\rho_w := \sum_{b\in B} w_b < 1/3$,*

then the weighted geometric median:

$$m = \arg\min_{y\in\mathbb{R}} \sum_{a\in\mathcal{A}} w_a \, |y - c_a|$$

satisfies:

$$|m - z| \le \sigma \left(1 + 4\sqrt{\frac{\rho_w}{1-\rho_w}}\right).$$

Proof. Define the honest centroid:

$$\bar{c}_H = \frac{\sum_{a\in H} w_a c_a}{\sum_{a\in H} w_a}.$$

Since $|c_a - z| \le \sigma$ for all $a \in H$,

$$|\bar{c}_H - z| \le \frac{\sum_{a\in H} w_a \, |c_a - z|}{\sum_{a\in H} w_a} \le \sigma.$$

By Theorem 2.8 (Minsker's bound) applied to the weighted setting:

$$|m - \bar{c}_H| \le 4\sigma\sqrt{\frac{\rho_w}{1-\rho_w}}.$$

By the triangle inequality,

$$|m - z| \le |m - \bar{c}_H| + |\bar{c}_H - z| \le 4\sigma\sqrt{\frac{\rho_w}{1-\rho_w}} + \sigma.$$

□

Corollary 4.2 (Mode-wise Application)**.** *For mode k, let honest agents satisfy $\left|c_k^{(a)} - c_k^\star\right| \le \sigma_k$ where $c_k^\star = \phi_k^\top f^\star$ is the ground truth coefficient. The geometric median m_k of $\{c_k^{(a)}\}_{a\in\mathcal{A}}$ satisfies:*

$$|m_k - c_k^\star| \le \sigma_k \left(1 + 4\sqrt{\frac{\rho_w}{1-\rho_w}}\right) =: C_{\text{med}}\sigma_k,$$

where $C_{\text{med}} = 1 + 4\sqrt{\rho_w/(1-\rho_w)}$.

4.2 Trimmed Weighted Mean

The geometric median provides robustness but requires iterative optimization (Weiszfeld's algorithm). We follow it with distance-based trimming to enable fast linear aggregation.

Theorem 4.3 (Trimmed Mean Bound)**.** *Under the setup of Theorem 4.1, let ρ_w denote the total Byzantine weight, and assume $\rho_w < 1/3$. Fix a trimming parameter α satisfying*

$$\rho_w \leq \alpha \leq 1 - 2\rho_w.$$

Given scalar values $\{c_a\}_{a\in\mathcal{A}}$ with weights $\{w_a\}_{a\in\mathcal{A}}$ such that $\sum_a w_a = 1$, compute:

1. *geometric median m_k;*
2. *distances $d_a = |c_a - m_k|$ for all $a \in \mathcal{A}$;*
3. *trim agents with largest d_a (breaking ties arbitrarily) until the total trimmed weight W_T satisfies $W_T \leq \alpha$ and adding any further agent would make $W_T > \alpha$;*
4. *let $S \subset \mathcal{A}$ be the remaining (kept) agents;*
5. *compute the trimmed mean with renormalized weights $\widetilde{w}_a = w_a / \sum_{j\in S} w_j$:*

$$\hat{c} = \sum_{a\in S} \widetilde{w}_a c_a.$$

Let $C_{\text{med}} = 1 + 4\sqrt{\rho_w/(1-\rho_w)}$ be the geometric median constant from Theorem 4.1, and define

$$R_H := \sigma(1 + C_{\text{med}}) = \sigma\left(2 + 4\sqrt{\frac{\rho_w}{1-\rho_w}}\right).$$

Then

$$|\hat{c} - z| \leq R_H.$$

Proof. Let H and B denote the sets of honest and Byzantine agents, respectively, with total Byzantine weight $\rho_w = \sum_{b\in B} w_b$.

Step 1: Honest agents are close to the median. By assumption, for each honest agent $h \in H$,

$$|c_h - z| \leq \sigma.$$

By Theorem 4.1, the geometric median m_k satisfies

$$|m_k - z| \leq C_{\text{med}}\sigma.$$

Hence for any honest $h \in H$,

$$d_h = |c_h - m_k| \leq |c_h - z| + |z - m_k| \leq \sigma + C_{\text{med}}\sigma = R_H.$$

Thus every honest agent lies in the interval $\{a : d_a \leq R_H\}$ around m_k.

Step 2: Trimming preserves an honest majority in weight. Let

$$W_H = \sum_{h \in H} w_h = 1 - \rho_w, \qquad W_T = \sum_{a \notin S} w_a \leq \alpha,$$

be the total honest and trimmed weights, respectively. In the worst case, all trimmed weight is honest, so the honest weight remaining in S satisfies

$$W_{S_H} := \sum_{a \in S \cap H} w_a \geq W_H - W_T \geq (1 - \rho_w) - \alpha.$$

Byzantine weight in S is at most

$$W_{S_B} := \sum_{a \in S \cap B} w_a \leq \rho_w.$$

Total kept weight is

$$W_S := \sum_{a \in S} w_a = 1 - W_T \geq 1 - \alpha.$$

Therefore the *honest weight fraction* after trimming satisfies

$$\widetilde{W}_{S_H} := \frac{W_{S_H}}{W_S} \geq \frac{1 - \rho_w - \alpha}{1 - \alpha}.$$

Using $\alpha \leq 1 - 2\rho_w$ and $\rho_w < 1/3$,

$$\frac{1 - \rho_w - \alpha}{1 - \alpha} \;\geq\; \frac{1 - \rho_w - (1 - 2\rho_w)}{1 - (1 - 2\rho_w)} \;=\; \frac{\rho_w}{2\rho_w} = \frac{1}{2}.$$

Hence

$$\widetilde{W}_{S_H} \geq \frac{1}{2}, \qquad \widetilde{W}_{S_B} := \frac{W_{S_B}}{W_S} = 1 - \widetilde{W}_{S_H} \leq \frac{1}{2}.$$

Step 3: Kept Byzantine agents cannot be too far from m_k. Define the "far" set

$$F := \{a \in \mathcal{A} : d_a > R_H\}.$$

By Step 1, all honest agents have $d_h \leq R_H$, so $F \subseteq B$ and its total weight satisfies

$$W_F := \sum_{a \in F} w_a \leq \rho_w.$$

The trimming procedure removes agents in order of decreasing d_a until the total trimmed weight reaches some $W_T \leq \alpha$, and by assumption $\alpha \geq \rho_w \geq W_F$. Thus the algorithm can remove all of F without exceeding the trimming budget α; since we always remove the farthest available agents first, every agent in F is trimmed. Hence

$$F \cap S = \emptyset,$$

and every kept agent $a \in S$ satisfies $d_a \leq R_H$.

For any kept Byzantine agent $b \in S \cap B$, we therefore have

$$|c_b - z| \le |c_b - m_k| + |m_k - z| \le R_H + C_{\mathrm{med}}\sigma = \sigma(1 + 2C_{\mathrm{med}}).$$

Step 4: Bound the trimmed mean. We can now bound the deviation of the trimmed mean:

$$\begin{aligned} |\hat{c} - z| &= \left| \sum_{a \in S} \widetilde{w}_a (c_a - z) \right| \\ &\le \sum_{a \in S_H} \widetilde{w}_a \, |c_a - z| + \sum_{a \in S_B} \widetilde{w}_a \, |c_a - z| \\ &\le \widetilde{W}_{S_H} \cdot \sigma + \widetilde{W}_{S_B} \cdot \sigma(1 + 2C_{\mathrm{med}}), \end{aligned}$$

where $\widetilde{W}_{S_H} = \sum_{a \in S_H} \widetilde{w}_a$ and similarly for $\widetilde{W}_{S_B}$. Using $\widetilde{W}_{S_H} + \widetilde{W}_{S_B} = 1$ and $\widetilde{W}_{S_B} \le 1/2$, we get

$$\begin{aligned} |\hat{c} - z| &\le \sigma + \widetilde{W}_{S_B} \, \sigma(2C_{\mathrm{med}}) \\ &\le \sigma + \frac{1}{2}\sigma(2C_{\mathrm{med}}) \\ &= \sigma(1 + C_{\mathrm{med}}) \\ &= R_H. \end{aligned}$$

This completes the proof. □

Remark 4.4. In the spectral protocol we use mode-dependent trimming levels α_k satisfying $\rho_w \le \alpha_k \le 1 - 2\rho_w$ (e.g. $\alpha_k = \mathrm{clip}(\eta\lambda_k, \rho_w, 1 - 2\rho_w)$ for a small stepsize η). This only changes constants, not the asymptotic dependence on ρ_w or σ_k.

Constants summary:

- $C_{\mathrm{med}} = 1 + 4\sqrt{\rho_w/(1 - \rho_w)}$ (geometric median);
- $C_{\mathrm{trim}} = 1 + C_{\mathrm{med}} = 2 + 4\sqrt{\rho_w/(1 - \rho_w)}$ (trimmed mean).

For $\rho_w = 0.3$ (near the maximum), $C_{\mathrm{trim}} \approx 4.62$.

5 Main Results II: Convergence Analysis

5.1 Gossip Contraction with Byzantine Noise

We model the gossip process on the honest subgraph G_H.

Assumption 5.1 (Honest Gossip Graph)**.** The honest agents H communicate over a connected graph $G_H = (H, E_H)$ with normalized Laplacian L_H. The gossip process is a lazy random walk on G_H with transition matrix:

$$P = I - \frac{1}{2}L_H$$

(lazy random walk with spectral gap $\lambda_\star := \lambda_1(L_H) > 0$).

Byzantine agents send values but do not appear in neighbor sampling (i.e., agents only sample neighbors from H).

Remark 5.2. We consider a standard Byzantine adversary that controls the values sent by agents in B, but cannot alter the topology or randomness of the honest gossip process on G_H. This aligns the model with the contraction properties of the lazy random walk on the honest subgraph.

Lemma 5.3 (Weighted Gossip Contraction)**.** *Under Assumption 5.1, for mode k at round t, let:*

- $c_{i,k}^{(t)}$*: agent i's current coefficient;*
- $\bar{c}_k^{(t)} = |H|^{-1} \sum_{i \in H} c_{i,k}^{(t)}$*: honest average;*
- $\tilde{c}_{i,k}^{(t)}$*: aggregated value from protocol (includes Byzantine contributions).*

Define the Byzantine noise:

$$\xi_{i,k}^{(t)} := \tilde{c}_{i,k}^{(t)} - \sum_{j \in H} P_{ij} c_{j,k}^{(t)}.$$

Then:

$$\mathbb{E}\big[\big|c_{i,k}^{(t)} - \bar{c}_k^{(t)}\big|^2\big] \le (1 - \lambda_\star)\big|c_{i,k}^{(t)} - \bar{c}_k^{(t)}\big|^2 + \mathbb{E}\big[\Big|\xi_{i,k}^{(t)}\Big|^2\big],$$

where expectation is over the gossip randomness.

Moreover, by Theorem 4.3,

$$\mathbb{E}\big[\Big|\xi_{i,k}^{(t)}\Big|^2\big] \le C_{\mathrm{trim}}^2 \sigma_k^2 \frac{\rho_w}{1 - \rho_w} =: \sigma_\xi^2.$$

Proof. Split the aggregation into honest gossip and Byzantine perturbation:

$$\tilde{c}_{i,k}^{(t)} = \sum_{j \in H} P_{ij} c_{j,k}^{(t)} + \xi_{i,k}^{(t)}.$$

For the honest gossip term, by the standard Boyd et al. (2006) analysis of lazy random walk:

$$\mathbb{E}_{\mathrm{gossip}} \left[\left(\sum_{j \in H} P_{ij} c_{j,k}^{(t)} - \bar{c}_k^{(t)} \right)^2 \right] \le (1 - \lambda_\star)\big(c_{i,k}^{(t)} - \bar{c}_k^{(t)}\big)^2.$$

For the Byzantine term, $\xi_{i,k}^{(t)}$ represents the deviation introduced by trimmed aggregation. By Theorem 4.3, if the true honest average in the sampled neighborhood is $\bar{c}_{\mathrm{local}}$, then

$$\Big|\xi_{i,k}^{(t)}\Big| \le C_{\mathrm{trim}} \sigma_k \sqrt{\frac{\rho_w}{1 - \rho_w}}.$$

Taking expectations and using independence of gossip randomness and Byzantine perturbations (Byzantine agents do not know which neighbors were sampled),

$$\mathbb{E}\big[\Big|\tilde{c}_{i,k}^{(t)} - \bar{c}_k^{(t)}\Big|^2\big] \le (1 - \lambda_\star)\big|c_{i,k}^{(t)} - \bar{c}_k^{(t)}\big|^2 + \sigma_\xi^2.$$

□

5.2 Heat Flow Convergence

Theorem 5.4 (Mode-wise Exponential Decay)**.** *Under Assumption 5.1 and with update rule:*

$$c_{i,k}^{(t+1)} = (1-\alpha_k)c_{i,k}^{(t)} + \alpha_k \tilde{c}_{i,k}^{(t)},$$

where $\alpha_k = \eta\lambda_k$ with $\eta \in (0, 1/\lambda_{\max})$, the expected squared error satisfies:

$$\mathbb{E}\big[\big|c_{i,k}^{(t)} - c_k^{\infty}\big|^2\big] \le e^{-2\eta\lambda_k t} D_0^2 + C_{\text{trim}}^2 \sigma_k^2 \frac{\rho_w}{1-\rho_w}\big(1 - e^{-\eta\lambda_k t}\big)^2,$$

where:

- $D_0 = \max_{i,j\in H} \left|c_{i,k}^{(0)} - c_{j,k}^{(0)}\right|$ *is the initial disagreement;*
- c_k^{∞} *is the limiting consensus value (Byzantine-perturbed equilibrium).*

In particular, the steady-state variance is bounded by:

$$\lim_{t\to\infty} \mathbb{E}\big[\big|c_{i,k}^{(t)} - c_k^{\star}\big|^2\big] \le C_{\text{trim}}^2 \sigma_k^2 \frac{\rho_w}{1-\rho_w},$$

where $c_k^{\star}$ is the ground truth coefficient.

Proof. Define the deviation from the honest average:

$$e_{i,k}^{(t)} := c_{i,k}^{(t)} - \bar{c}_k^{(t)}.$$

From the update rule,

$$\begin{aligned} e_{i,k}^{(t+1)} &= (1-\alpha_k)e_{i,k}^{(t)} + \alpha_k\big(\tilde{c}_{i,k}^{(t)} - \bar{c}_k^{(t)}\big) \\ &= (1-\alpha_k)e_{i,k}^{(t)} + \alpha_k\left(\sum_{j\in H} P_{ij}(c_{j,k}^{(t)} - \bar{c}_k^{(t)}) + \xi_{i,k}^{(t)}\right). \end{aligned}$$

By Lemma 5.3,

$$\mathbb{E}\Big[\Big|e_{i,k}^{(t+1)}\Big|^2\Big] \le (1-\alpha_k)^2 \Big|e_{i,k}^{(t)}\Big|^2 + \alpha_k^2\big[(1-\lambda_\star)\Big|e_{i,k}^{(t)}\Big|^2 + \sigma_\xi^2\big].$$

Simplifying,

$$\begin{aligned} \mathbb{E}\Big[\Big|e_{i,k}^{(t+1)}\Big|^2\Big] &\le \big[(1-\alpha_k)^2 + \alpha_k^2(1-\lambda_\star)\big]\Big|e_{i,k}^{(t)}\Big|^2 + \alpha_k^2\sigma_\xi^2 \\ &= \big[1 - 2\alpha_k + \alpha_k^2 + \alpha_k^2(1-\lambda_\star)\big]\Big|e_{i,k}^{(t)}\Big|^2 + \alpha_k^2\sigma_\xi^2 \\ &= \big[1 - 2\alpha_k + \alpha_k^2(2-\lambda_\star)\big]\Big|e_{i,k}^{(t)}\Big|^2 + \alpha_k^2\sigma_\xi^2. \end{aligned}$$

For small $\alpha_k = \eta\lambda_k \ll 1$, the contraction coefficient is approximately

$$1 - 2\alpha_k \approx e^{-2\alpha_k} = e^{-2\eta\lambda_k}.$$

Unrolling the recursion,

$$\mathbb{E}\big[\big|e_{i,k}^{(t)}\big|^2\big] \le e^{-2\eta\lambda_k t}\left|e_{i,k}^{(0)}\right|^2 + \sigma_\xi^2 \sum_{s=0}^{t-1} \alpha_k^2 e^{-2\eta\lambda_k(t-s)}$$
$$\le e^{-2\eta\lambda_k t} D_0^2 + \sigma_\xi^2 \alpha_k^2 \frac{1-e^{-2\eta\lambda_k t}}{1-e^{-2\eta\lambda_k}}.$$

For small α_k, we have

$$\frac{\alpha_k^2}{1-e^{-2\eta\lambda_k}} = O(\alpha_k),$$

and we obtain the stated bound (up to a mild constant relabeling) with the noise term behaving like $\sigma_\xi^2(1-e^{-\eta\lambda_k t})^2$.

Substituting $\sigma_\xi^2 = C_{\text{trim}}^2\sigma_k^2\rho_w/(1-\rho_w)$ completes the proof. □

5.3 Anisotropic Lyapunov Contraction

We now establish global convergence by defining a weighted variance functional.

Definition 5.5 (Weighted Variance Lyapunov Function)**.** Define the per-mode variance among honest agents:

$$V_k^{(t)} := \frac{1}{|H|}\sum_{i\in H}\big(c_{i,k}^{(t)} - \bar{c}_k^{(t)}\big)^2,$$

and the weighted total variance:

$$V^{(t)} := \sum_{k=0}^{K-1} w_k V_k^{(t)},$$

where $w_k = (1+\lambda_k)^{-\beta}$ for some $\beta > 0$ (mode weighting exponent).

Theorem 5.6 (Global Lyapunov Contraction)**.** *Under the heat flow protocol with $\alpha_k = \eta\lambda_k e^{-k/K}$ (exponentially decreasing damping), the weighted variance satisfies:*

$$\mathbb{E}\big[V^{(t+1)}\big] \le (1-\gamma)V^{(t)} + \Delta_B,$$

where:

- *contraction coefficient $\gamma \approx \alpha_0\lambda_\star/2$ with $\alpha_0 = \eta\lambda_0 \approx 0$ (smallest mode);*
- *Byzantine noise floor $\Delta_B = C_{\text{trim}}^2 \sum_{k=0}^{K-1} w_k\sigma_k^2 \frac{\rho_w}{1-\rho_w}$.*

Unrolling,

$$V^{(t)} \le e^{-\gamma t}V^{(0)} + \frac{\Delta_B}{\gamma}\big(1-e^{-\gamma t}\big).$$

Convergence time to ϵ-consensus is

$$T_\epsilon = O\left((\alpha_0\lambda_\star)^{-1}\log\frac{V^{(0)}}{\epsilon}\right),$$

where $V^{(0)}$ depends on initial disagreement D_0^2 across honest agents.

Proof. From Lemma 5.3, for each mode k,

$$\mathbb{E}\big[V_k^{(t+1)}\big] \le (1 - 2\alpha_k\lambda_\star)V_k^{(t)} + \alpha_k^2\sigma_\xi^2.$$

Multiplying by w_k and summing over k,

$$\begin{aligned}\mathbb{E}\big[V^{(t+1)}\big] &= \sum_{k=0}^{K-1} w_k \mathbb{E}\big[V_k^{(t+1)}\big] \\ &\le \sum_{k=0}^{K-1} w_k(1 - 2\alpha_k\lambda_\star)V_k^{(t)} + \sum_{k=0}^{K-1} w_k\alpha_k^2\sigma_{\xi,k}^2 \\ &= \sum_{k=0}^{K-1} w_k V_k^{(t)} - 2\lambda_\star \sum_{k=0}^{K-1} w_k\alpha_k V_k^{(t)} + \Delta_B.\end{aligned}$$

Lower-bound the contraction term:

$$\sum_{k=0}^{K-1} w_k\alpha_k V_k^{(t)} \ge \min_k(\alpha_k) \sum_{k=0}^{K-1} w_k V_k^{(t)} = \alpha_0 e^{-(K-1)/K} V^{(t)} \approx 0.37\,\alpha_0 V^{(t)}.$$

Therefore,

$$\mathbb{E}\big[V^{(t+1)}\big] \le \big(1 - 2 \cdot 0.37\lambda_\star\alpha_0\big)V^{(t)} + \Delta_B \approx (1 - 0.74\lambda_\star\alpha_0)V^{(t)} + \Delta_B.$$

Setting $\gamma = \alpha_0\lambda_\star/2$ (conservative) gives the stated bound. The unrolling and convergence time follow from standard AR(1) analysis. □

Corollary 5.7 (Mode-Dependent Convergence Rates)**.** *For $\rho_w < 1/3$ and $\sigma_k^2 = \sigma_0^2/\lambda_k$ (natural noise scaling),*

$$\Delta_B = C_{\text{trim}}^2\sigma_0^2 \frac{\rho_w}{1 - \rho_w} \sum_{k=0}^{K-1} \frac{w_k}{\lambda_k}.$$

For $w_k = (1 + \lambda_k)^{-\beta}$ and uniform $\lambda_k \approx k/K$,

$$\Delta_B \approx C_{\text{trim}}^2\sigma_0^2 \frac{\rho_w}{1 - \rho_w} \cdot K \cdot \log K,$$

i.e., logarithmic dependence on K.

6 Main Results III: Dynamic Topology and Fork Detection

6.1 Time-Varying Honest Graphs

Assumption 6.1 (Uniform Spectral Gap)**.** The honest communication graph $G_H^{(t)}$ may change over time, but satisfies:

$$\lambda_1(L_H^{(t)}) \ge \lambda_\star > 0 \quad \forall t,$$

a uniform lower bound on spectral gap.

Theorem 6.2 (Dynamic Topology Resilience)**.** *Under Assumption 6.1, Theorems 5.4 and 5.6 hold with spectral gap replaced by $\lambda_\star$ (the worst-case gap). In particular, the convergence time becomes*

$$T_\epsilon = O\left(\frac{1}{\alpha_0 \lambda_\star} \log \frac{1}{\epsilon}\right),$$

where the dependence on $\lambda_\star$ reflects the "bottleneck" epoch with poorest connectivity.

Proof. At each round t, apply Lemma 5.3 with the current graph's spectral gap $\lambda_1(L_H^{(t)})$. The variance contraction is:

$$\mathbb{E}\big[V_k^{(t+1)} \mid G_H^{(t)}\big] \leq (1 - 2\alpha_k \lambda_1(L_H^{(t)}))V_k^{(t)} + \sigma_\xi^2.$$

Taking expectations over the sequence of graphs and using $\lambda_1(L_H^{(t)}) \geq \lambda_\star$,

$$\mathbb{E}\big[V_k^{(t+1)}\big] \leq (1 - 2\alpha_k \lambda_\star)V_k^{(t)} + \sigma_\xi^2.$$

The rest of the proof proceeds identically to Theorem 5.6. □

6.2 Spectral Fork Detection and GeoCerts

Assumption 6.3 (Dirichlet Energy Constraint)**.** The fork detection mechanism operates under the assumption that valid events satisfy:

$$\mathcal{E}(e) = f_e^\top L f_e = \sum_{k=0}^{n-1} \lambda_k c_k(e)^2 \leq \mathcal{E}_0 + \tau.$$

Enforcement of this constraint requires external mechanisms (proof-of-work, stake slashing, or application-layer validation). Theorem 6.5 characterizes what becomes detectable when this constraint holds.

Definition 6.4 (Spectral Coherence Certificate (GeoCert))**.** Given an event e with spectral coefficients $\{c_k(e)\}_{k=0}^{K-1}$ and Dirichlet energy bound $\mathcal{E}(e) \leq \mathcal{E}_0 + \tau$, a *spectral coherence certificate* (or *GeoCert*) for e is the tuple

$$\mathsf{GeoCert}(e) := \big(K,\ \{\lambda_k\}_{k=0}^{K-1},\ \{c_k(e)\}_{k=0}^{K-1},\ \mathcal{E}(e),\ \tau\big)$$

together with the verification rule of Theorem 6.5: any triple (e_0, e_1, e_2) whose truncated spectral distances violate the triangle inequality by more than the prescribed margin is flagged as geometrically inconsistent.

Theorem 6.5 (Spectral Fork Detection (under Assumption 6.3))**.** *Let e_0, e_1, e_2 be three events with spectral coefficients truncated to K modes. Define the weighted spectral distance:*

$$d_\Lambda(e_i, e_j) := \left(\sum_{k=0}^{K-1} (1 + \lambda_k)^{-\beta} \, |c_k(e_i) - c_k(e_j)|^2\right)^{1/2}.$$

Under Assumption 6.3, if

$$d_\Lambda(e_1, e_0) + d_\Lambda(e_2, e_0) \leq d_\Lambda(e_1, e_2) - \epsilon$$

for some

$$\epsilon > C_{\text{fork}} \sqrt{\frac{\tau}{\lambda_K^{1+\beta}}}$$

where $C_{\text{fork}} = 3$ is a universal constant, then at least one of $\{e_1, e_2\}$ violates the energy budget (i.e., has $\mathcal{E}(e_i) > \mathcal{E}_0 + \tau$).

Proof. **Step 1: Bound high-frequency mass.** From Assumption 6.3,

$$\sum_{k=K}^{n-1} \lambda_k c_k(e)^2 \leq \mathcal{E}(e) - \sum_{k=0}^{K-1} \lambda_k c_k(e)^2 \leq \mathcal{E}_0 + \tau.$$

For $k \geq K$, $\lambda_k \geq \lambda_K$, so

$$\sum_{k=K}^{n-1} c_k(e)^2 \leq \frac{\mathcal{E}_0 + \tau}{\lambda_K} =: E_{\text{high}}.$$

Step 2: Weighted norm truncation error. The truncation error in weighted norm is

$$\begin{aligned} \|c(e) - c_{\text{trunc}}(e)\|_\Lambda^2 &= \sum_{k=K}^{n-1} (1 + \lambda_k)^{-\beta} c_k(e)^2 \\ &\leq (1 + \lambda_K)^{-\beta} \sum_{k=K}^{n-1} c_k(e)^2 \\ &\leq \frac{E_{\text{high}}}{(1 + \lambda_K)^\beta} \leq \frac{\tau}{\lambda_K^{1+\beta}}, \end{aligned}$$

assuming $\mathcal{E}_0 \ll \tau$ for simplicity; otherwise add a constant.

Step 3: Triangle inequality in full space. In the full n-dimensional spectral space, triangle inequality must hold:

$$\|c(e_1) - c(e_0)\|_\Lambda + \|c(e_2) - c(e_0)\|_\Lambda \geq \|c(e_1) - c(e_2)\|_\Lambda .$$

Step 4: Relate truncated to full distances. For any two events e_i, e_j,

$$\begin{aligned} \|c(e_i) - c(e_j)\|_\Lambda &\leq \|c_{\text{trunc}}(e_i) - c_{\text{trunc}}(e_j)\|_\Lambda + \|c(e_i) - c_{\text{trunc}}(e_i)\|_\Lambda + \|c(e_j) - c_{\text{trunc}}(e_j)\|_\Lambda \\ &\leq d_\Lambda(e_i, e_j) + 2\sqrt{\frac{\tau}{\lambda_K^{1+\beta}}}. \end{aligned}$$

Similarly,

$$\|c(e_i) - c(e_j)\|_\Lambda \geq d_\Lambda(e_i, e_j) - 2\sqrt{\frac{\tau}{\lambda_K^{1+\beta}}}.$$

Step 5: Combine. Assume the hypotheses:

$$d_\Lambda(e_1, e_0) + d_\Lambda(e_2, e_0) \leq d_\Lambda(e_1, e_2) - \epsilon.$$

In the full space,

$$\begin{aligned}
&\|c(e_1) - c(e_0)\|_\Lambda + \|c(e_2) - c(e_0)\|_\Lambda \\
&\quad \le d_\Lambda(e_1, e_0) + 2\sqrt{\frac{\tau}{\lambda_K^{1+\beta}}} + d_\Lambda(e_2, e_0) + 2\sqrt{\frac{\tau}{\lambda_K^{1+\beta}}} \\
&\quad \le d_\Lambda(e_1, e_2) - \epsilon + 4\sqrt{\frac{\tau}{\lambda_K^{1+\beta}}} \\
&\quad < \|c(e_1) - c(e_2)\|_\Lambda
\end{aligned}$$

if $\epsilon > 6\sqrt{\tau/\lambda_K^{1+\beta}}$ (using the lower bound on $\|c(e_1) - c(e_2)\|_\Lambda$). This violates triangle inequality in the full space, which is only possible if one of e_1, e_2 violates the energy constraint. Taking $C_{\text{fork}} = 3$ (accounting for the factor of 2 in each direction) gives the stated threshold. □

Interpretation: If two proposed events e_1, e_2 are both claimed to be valid continuations from e_0, but their truncated spectral representations form a "nearly degenerate triangle" (sum of two sides $\approx$ third side minus ϵ), this is geometrically impossible unless one event spent more energy than allowed. This provides a *computable geometric certificate* of inconsistency, though practical deployment requires careful threshold calibration and energy budget enforcement mechanisms.

Table 1: Comparison with Byzantine-Robust Aggregation Methods

Method	**Per-Round Cost**	ρ **Tolerance**	**Geometry-Aware**
Krum [1]	$O(n^2 d)$	$< 1/2$	No
Geometric Median [7]	$O(nd \log(1/\epsilon))$	$< 1/3$	No
Bulyan [6]	$O(n^2 d)$	$< 1/3$	No
This work	$O(K\|E\| + Kn)$	$\rho_w < 1/3$	**Yes**

7 Limitations and Practical Considerations

7.1 Energy Budget Enforcement

Critical caveat: Theorem 6.5 provides a *detection mechanism* but not autonomous enforcement. The protocol cannot prevent adversaries from violating the energy bound $\mathcal{E}(e) \le \mathcal{E}_0 + \tau$ without external mechanisms such as:

- **Proof-of-work**: Computational puzzles proportional to Dirichlet energy, making high-energy events expensive to produce;
- **Stake slashing**: Economic penalties for agents whose proposed events are geometrically detected as violating energy constraints;
- **Application-layer validation**: Domain-specific bounds (e.g., physical plausibility constraints in multi-robot SLAM).

Without such enforcement, Theorem 6.5 identifies *which* events are inconsistent but cannot force adversaries to respect energy limits. This represents a fundamental tradeoff: geometric impossibility provides *certificates* of misbehavior, but economic or cryptographic mechanisms are needed to make such misbehavior *costly*.

7.2 Manifold Learning Under Adversarial Conditions

The protocol assumes honest agents share a common learned Laplacian L. If Byzantine agents control the manifold learning phase (e.g., by poisoning training data used to construct embeddings), they may craft pathological eigenspaces where:

- High-frequency adversarial perturbations masquerade as low-frequency structure (misclassified smoothness);
- Spectral coherence detection degrades or fails entirely.

Mitigation strategies:

- **Robust manifold learning**: Use techniques such as robust PCA or spectral clustering algorithms that resist poisoning;
- **Trusted initialization**: Bootstrap from a secure setup phase where L is computed before Byzantine agents join;
- **Periodic re-learning**: Refresh L using only recent data from high-reputation agents, limiting adversarial accumulation.

This remains an open research problem; existing manifold learning methods (Isomap, Laplacian Eigenmaps, UMAP) lack Byzantine robustness guarantees.

7.3 Scalability and Communication Overhead

While the per-round complexity $O(K|E| + KN \log N)$ is polynomial, practical deployment faces bottlenecks:

- **Spectral projection cost**: Computing K eigenvectors via Lanczos or Chebyshev approximation requires $O(K|E|)$ operations, which becomes prohibitive for dense graphs ($|E| = \Theta(n^2)$) or large K;
- **Coefficient broadcast**: Each agent transmits K coefficients per round. For $K = 100$ and float32 precision, this is $\sim$400 bytes/agent, manageable for $N \leq 10^3$ but challenging for planetary-scale networks ($N \sim 10^6$);
- **Geometric median iterations**: Weiszfeld's algorithm requires $O(\log(1/\epsilon))$ iterations. With m neighbors, this is $O(m \log(1/\epsilon))$ per mode, dominating runtime for large neighborhoods.

Complexity analysis suggests practical deployment at $N \approx 10^2$–10^3 agents and $K \approx 20$–100 modes, with consensus time $O(\text{seconds})$ on typical networks.

7.4 Applications and Non-Applications

Where this protocol excels:

- **Federated learning on graphs**: Aggregating gradients or model weights for GNNs, molecular property prediction, or knowledge graph embeddings where data has natural graph structure;
- **Multi-agent semantic SLAM**: Robots building shared 3D scene representations where consensus on low-frequency geometry (walls, floors) is critical but high-frequency detail (texture) can vary;
- **Narrative consensus engines**: AI agents in games/simulations agreeing on "canon" story elements (global state) while preserving local variations (character perspectives).

Where traditional BFT is superior:

- **High-throughput payment systems**: Transaction ordering requires millisecond finality at 10^4–10^6 TPS, far beyond this protocol's 1–10 second consensus time;
- **Discrete event logs**: Blockchains, audit trails, and databases where semantic structure is absent and total ordering is paramount;
- **Low-bandwidth environments**: Networks with <1 Mbps connectivity cannot afford K-dimensional coefficient broadcasts.

The protocol targets *semantic* consensus (1–10s latency acceptable) on *geometric semantic data*, not discrete transaction sequencing.

8 Related Work

8.1 Byzantine Fault Tolerance

Classical BFT (Castro & Liskov 1999 [4], Buchman 2016 [3]): State machine replication with $3f + 1$ replicas to tolerate f Byzantine faults. Achieves consensus on discrete transaction order via multi-round voting ($O(n^2)$ messages).

HotStuff (Yin et al. 2019 [10]): Linear communication ($O(n)$) via leader-based pipelining and threshold signatures. Optimizes latency for permissioned blockchains.

Gap: these protocols operate on *discrete state* (transaction sets, block hashes), while our work addresses *continuous semantic fields* with geometric structure.

8.2 Byzantine-Robust Aggregation

Krum (Blanchard et al. 2017 [1]): Byzantine-robust gradient aggregation in federated learning. Selects gradient with smallest sum of distances to neighbors. Time complexity $O(n^2 d)$.

Geometric Median (Minsker 2015 [7]): Proves $\|m - \mu\|_2 \leq C\sigma\sqrt{f/n}$ for f Byzantine agents. Optimal rate but requires iterative optimization.

Bulyan (Mhamdi et al. 2018 [6]): Combines Krum with trimmed mean for label-flipping resilience.

Our contribution: we extend geometric median to *spectral space* with mode-dependent weighting, achieving anisotropic robustness where low-frequency (global) modes are prioritized.

8.3 Spectral Methods in Consensus

Gossip Algorithms (Boyd et al. 2006 [2]): Distributed averaging via pairwise mixing. Convergence rate $O(\lambda_1^{-1} \log(1/\epsilon))$ governed by spectral gap.

Graph Signal Processing (Shuman et al. 2013 [9]): Defines graph Fourier transform via Laplacian eigenvectors. Our Chebyshev projection is a fast approximation.

8.4 Geometric Methods in Distributed Systems

Riemannian Consensus (Sarlette et al. 2009 [8]): Consensus on manifolds (SO(3), Grassmann) via Fréchet mean. Does not address Byzantine faults.

Wasserstein Barycenters (Cuturi & Doucet 2014 [5]): Optimal transport-based averaging of distributions. Computationally expensive ($O(n^3)$ via Sinkhorn).

The protocol operates in spectral domain with fast Chebyshev approximation, Byzantine robustness via geometric median, and anisotropic damping.

8.5 Positioning

Our protocol combines spectral representation, Byzantine-robust aggregation via geometric median, and anisotropic convergence for distributed AI applications (federated learning on graphs, multi-agent world modeling). This represents an unexplored point in the design space between classical BFT (discrete state, voting-based) and distributed machine learning (continuous parameters, non-adversarial averaging).

9 Future Directions

9.1 Open Problems

Optimal mode weighting: characterize the optimal choice of β in $w_k = (1 + \lambda_k)^{-\beta}$ for different graph structures and noise models.

Adaptive trimming thresholds: current analysis assumes fixed α for trimming. Adaptive selection based on observed Byzantine fraction could improve efficiency.

Sybil attack analysis: investigate whether splitting Byzantine stake into multiple identities reduces or increases attack effectiveness under geometric median aggregation.

9.2 Extensions

Adaptive mode selection: dynamically choose K based on observed noise levels and computational constraints.

Higher-order Laplacians: extend to simplicial complexes for modeling relational data (hypergraphs, knowledge graphs).

Nonlinear manifolds: incorporate Riemannian metric structure for consensus on curved spaces (SO(3) for poses, probability simplex for distributions).

Remark 9.1. Economic Enforcement and Perceptual Architectures). Theorem 6.5 provides geometric fork detection but requires external enforcement of an energy budget. A natural extension is to integrate game-theoretic mechanisms (e.g., stake slashing or reputation-weighted aggregation) so that geometric violations are economically irrational. Ongoing work at Peak&Tower explores such couplings, but formal incentive-compatibility results are left for future work.

Hypergossip and dynamic geometry. In this work the geometric backbone (graph G and Laplacian L) is fixed. A natural extension is to let the semantic manifold evolve jointly with consensus state, for example by learning a data-driven Laplacian from agent embeddings and coupling it to conviction or reputation variables. This leads to "hypergossip" architectures where consensus dynamics deform the geometry itself, enabling adaptation to non-stationary semantic spaces.

10 Conclusion

We have presented a mathematical framework for Byzantine-robust consensus on continuous semantic fields, bridging robust statistics, spectral graph theory, and distributed systems. Our key innovations—mode-wise geometric median aggregation, anisotropic heat flow dynamics, and spectral coherence detection—enable provably secure consensus on high-dimensional geometric data with applications to distributed AI and federated learning on graph-structured domains.

The framework provides:

- **Theoretical guarantees**: per-mode error bounds $C_{\text{trim}}\sigma_k\sqrt{\rho_w/(1-\rho_w)}$ for Byzantine fraction $\rho_w < 1/3$ (Theorems 4.1, 4.3);
- **Convergence analysis**: exponential contraction of weighted variance with rate $\gamma \approx \alpha_0\lambda_\star/2$ and convergence time $O((\alpha_0\lambda_\star)^{-1}\log(1/\epsilon))$ (Theorems 5.4, 5.6);
- **Scalable algorithms**: Chebyshev approximation achieving $O(K|E|)$ complexity (Lemma 2.5);
- **Geometric detection**: computable fork certificates via energy-constrained triangle inequality violations (Theorem 6.5), subject to external energy budget enforcement.

This work opens new directions at the intersection of Byzantine fault tolerance, geometric machine learning, and spectral methods, with immediate applications to federated learning on graphs, multi-agent world modeling, and decentralized AI inference systems. Key open problems include optimal mode weighting, adaptive Byzantine detection thresholds, and extension to nonlinear manifold geometries and dynamically evolving hypergossip-style metric deformations.

References

[1] P. Blanchard, E. M. El Mhamdi, R. Guerraoui, and J. Stainer. Machine learning with adversaries: Byzantine tolerant gradient descent. In *Advances in Neural Information Processing Systems*, 2017.

[2] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Randomized gossip algorithms. *IEEE Transactions on Information Theory*, 52(6):2508–2530, 2006.

[3] E. Buchman. Tendermint: Byzantine fault tolerance in the age of blockchains. Master's thesis, University of Guelph, 2016.

[4] M. Castro and B. Liskov. Practical byzantine fault tolerance. In *OSDI*, volume 99, pages 173–186, 1999.

[5] M. Cuturi and A. Doucet. Fast computation of wasserstein barycenters. In *International Conference on Machine Learning*, pages 685–693, 2014.

[6] E. M. El Mhamdi, R. Guerraoui, and S. Rouault. The hidden vulnerability of distributed learning in byzantium. In *International Conference on Machine Learning*, pages 3521–3530, 2018.

[7] S. Minsker. Geometric median and robust estimation in banach spaces. *Bernoulli*, 21(4):2308–2335, 2015.

[8] A. Sarlette, R. Sepulchre, and N. E. Leonard. Autonomous rigid body attitude synchronization. *Automatica*, 45(2):572–577, 2009.

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

[10] M. Yin, D. Malkhi, M. K. Reiter, G. G. Gueta, and I. Abraham. Hotstuff: Bft consensus with linearity and responsiveness. In *ACM Symposium on Principles of Distributed Computing*, pages 347–356, 2019.