Spectral of Spectrals: Second-Order Mode Decomposition for Complex Systems

Second-Order Mode Decomposition for Complex Systems

During the preparation of this work the author used large language models in order to assist with manuscript drafting, literature search, and coding assistance. After using these tools, the author reviewed and edited the content as needed and takes full responsibility for the content of the published article.

Tamás Nagy, Ph.D.

`tnagyphd@gmail.com`

Draft • March 2026

Abstract

Spectral decomposition extracts modes from data: eigenvectors of a covariance or kernel operator that capture maximal variance per component. We introduce a **second-order spectral layer**: the eigenanalysis of the spectral features themselves. Given a system where each point x has d first-order spectral diagnostics $\phi(x) = (\phi_1, \dots, \phi_d)$, we construct the feature covariance operator and extract its eigenmodes — the **meta-modes**. We prove that the meta-spectrum (1) is a sufficient statistic for second-order feature interactions, (2) achieves the optimal low-rank approximation of the feature correlation structure (Eckart-Young), and (3) provides a built-in significance test via comparison to the Marchenko-Pastur null distribution. On the Mandelbrot iteration, where first-level features include Lyapunov exponents, spectral radius, and derivative phase coherence, the meta-spectrum reveals that just **two meta-modes capture 92.6% of cross-feature variance**, providing a single decision surface that cleanly separates dynamical regimes. We describe extensions to kernel meta-spectra (nonlinear second-order analysis) and time-varying meta-spectra (regime tracking), and outline applications to agent routing, memory retrieval, proof search, and financial risk.

---

1. Introduction

1.1 The Curse of Multiple Diagnostics

Complex systems produce many spectral diagnostics: Lyapunov exponents, spectral gaps, coherence scores, condition numbers, energy distributions. Each diagnostic tells a partial story. The practitioner faces a combinatorial problem: which diagnostics to combine, how to weight them, and when to trust one over another.

The standard response is ad-hoc: pick a few diagnostics, threshold them, build a decision tree. This ignores correlations between diagnostics and misses structures that only appear in combinations.

1.2 The Meta-Spectral Idea

The insight is simple: **apply spectral analysis to the spectral features**. If you have d diagnostics measured at N points, you have a d-dimensional feature space. The eigendecomposition of its covariance reveals:

- How many **independent feature combinations** matter (effective dimensionality).
- Which **combinations** are dominant (meta-mode loadings).
- Where the **regime boundaries** lie (meta-mode sign changes).
- Which features are **redundant** (low meta-eigenvalue, high correlation with dominant modes).

This is not just PCA on features. The key difference is that the input features are themselves spectral — they are eigenvalue-derived quantities from a first-order analysis. The second-order structure captures **interactions between modes of the original system**.

---

2. Formal Framework

2.1 First-Order Spectral Features

Let $\mathcal{X}$ be a system (dynamical map, stochastic process, graph, proof library). At each point $x \in \mathcal{X}$, a first-order spectral analysis produces features:

$$\phi(x) = (\phi_1(x), \dots, \phi_d(x)) \in \mathbb{R}^d$$

Examples:

System	First-order features
Iterated map $z \mapsto z^2 + c$	Lyapunov exponent, spectral radius, phase coherence, escape time
Stochastic process $dX = \mu\, dt + \sigma\, dW$	Spectral gap, condition number, mixing time, stationary mode amplitudes
Knowledge graph	Spectral centrality, Fiedler value, effective resistance, bridge score
Agent state	Mode energies, gate values, uncertainty levels, policy confidence

2.2 Feature Covariance Operator

Given N observations $\{x_1, \dots, x_N\}$ with features $\Phi = [\phi(x_1), \dots, \phi(x_N)]^\top \in \mathbb{R}^{N\times d}$:

1. **Standardize**: $\tilde{\Phi}_{ij} = (\Phi_{ij} - \bar{\Phi}_j)/s_j$ where $\bar{\Phi}_j$ and s_j are the column mean and standard deviation.
2. **Covariance**: $C = \frac{1}{N-1}\tilde{\Phi}^\top\tilde{\Phi} \in \mathbb{R}^{d\times d}$.
3. **Eigendecompose**: $C = V\Lambda V^\top$ with $\Lambda = \mathrm{diag}(\lambda_1 \geq \dots \geq \lambda_d)$.

2.3 Meta-Modes

The j-th meta-mode is the projection of the standardized features onto the j-th eigenvector:

$$m_j(x_i) = \tilde{\phi}(x_i)^\top v_j$$

This gives a spatial field $m_j : \mathcal{X} \to \mathbb{R}$ for each meta-mode.

2.4 Key Quantities

Definition (Lead Share). $\ell_1 = \lambda_1 / \sum_j \lambda_j$. Measures how concentrated the meta-spectrum is. $\ell_1 \to 1$ means one meta-mode dominates.

Definition (Top-k Share). $\ell_k = \sum_{j=1}^{k} \lambda_j / \sum_j \lambda_j$. Fraction of feature-interaction variance captured by the first k meta-modes.

Definition (Meta-Resonance). $r(x) = |m_1(x)| / (|m_1(x)| + |m_2(x)| + \varepsilon) \in [0, 1]$. Measures how strongly the leading meta-mode dominates at point x. High r = the point's behavior is well-explained by a single feature combination. Low r = the point sits in a transition zone between meta-modes.

Definition (Effective Meta-Dimensionality). $d_{\text{eff}} = \left(\sum_j \lambda_j\right)^2 / \sum_j \lambda_j^2$. The participation ratio of the meta-spectrum. $d_{\text{eff}} = 1$ means one mode dominates; $d_{\text{eff}} = d$ means all modes are equally important.

---

3. Theoretical Properties

3.1 Optimality (Eckart-Young)

Theorem 1. The rank-k meta-mode approximation $\hat{\Phi}_k = \tilde{\Phi} V_k V_k^\top$ minimizes the Frobenius reconstruction error:

$$\|\tilde{\Phi} - \hat{\Phi}_k\|_F^2 = \sum_{j=k+1}^{d} \lambda_j$$

No other rank-k representation of the feature matrix achieves lower error.

3.2 Sufficient Statistic

Theorem 2. The meta-mode projections $(m_1, \ldots, m_k)$ are a sufficient statistic for all second-order interactions among the features $(\phi_1, \ldots, \phi_d)$, in the sense that:

$$\mathrm{Cov}(\phi_i, \phi_j) = \sum_{l=1}^{d} \lambda_l \, v_l^{(i)} \, v_l^{(j)}$$

Truncating to k meta-modes gives:

$$\mathrm{Cov}(\phi_i, \phi_j) \approx \sum_{l=1}^{k} \lambda_l \, v_l^{(i)} \, v_l^{(j)}$$

with error bounded by $\sum_{l=k+1}^{d} \lambda_l$.

3.3 Significance Testing (Marchenko-Pastur)

Theorem 3. Under the null hypothesis that the d features are independent with unit variance, the eigenvalues of C follow the Marchenko-Pastur distribution with support $[\lambda_-, \lambda_+]$ where:

$$\lambda_\pm = \left(1 \pm \sqrt{d/N}\right)^2$$

A meta-eigenvalue $\lambda_j > \lambda_+$ is significant at the α level determined by the Tracy-Widom distribution.

Practical test: For the Mandelbrot example ($d = 4$, $N = 48{,}400$), $\lambda_+ \approx 1.03$. The observed $\lambda_1 = 2.87$ is massively significant; $\lambda_2 = 0.84$ is marginally below the threshold but above the null mean, suggesting a weak but real second meta-mode.

* * *

4. Extensions

4.1 Kernel Meta-Spectrum (Nonlinear Second-Order Analysis)

If feature interactions are nonlinear, replace the linear covariance with a kernel operator:

$$\tilde{C}_{ij} = \frac{1}{N} \sum_{n=1}^{N} k(\phi_i(x_n), \phi_j(x_n))$$

where k is a kernel function (RBF, polynomial). The eigendecomposition of $\tilde{C}$ gives nonlinear meta-modes that capture curved interaction surfaces in feature space.

4.2 Time-Varying Meta-Spectrum

For time-series systems, compute the meta-spectrum on sliding windows:

$$C_t = \frac{1}{W-1} \tilde{\Phi}_{[t-W:t]}^\top \tilde{\Phi}_{[t-W:t]}$$

Track:

- **Lead share drift**: $\ell_1(t)$ over time. A drop in lead share signals increasing complexity (more meta-modes becoming relevant).
- **Meta-mode rotation**: the angle between consecutive leading eigenvectors $v_1(t)$ and $v_1(t-1)$. Large rotation = regime change.
- **Effective dimensionality drift**: $d_{\text{eff}}(t)$ tracks how many independent feature patterns are active.

4.3 Hierarchical Meta-Spectra

The meta-spectrum is itself a feature vector $(\lambda_1, \ldots, \lambda_d, v_1, \ldots, v_d)$. A third-order analysis — the spectrum of meta-spectra across different systems or different parameter regimes — could reveal **universal structural patterns** in how spectral features interact. This is speculative but connects to random matrix universality.

---

5. Demonstrated Example: Mandelbrot Iteration

5.1 Setup

System: $z_{n+1} = z_n^2 + c$ on a 220×220 grid over $[-2.1, 0.9] \times [-1.3, 1.3]$.

First-level features per c:

Feature	Symbol	Description
Lyapunov exponent	λ	$\log \rho(J_n)/n$
Log spectral radius	$\log \rho$	$\log(\max \lvert \mathrm{eig}(J_n) \rvert)$
Phase coherence	γ	Mean phase-locking of local derivatives
Normalized escape time	τ	Escape iteration / max iterations

5.2 Meta-Spectrum Results

Meta-mode	Eigenvalue	Share	Cumulative
1	2.866	71.65%	71.65%
2	0.845	21.12%	92.77%
3	0.213	5.32%	98.09%
4	0.076	1.91%	100.00%

Interpretation of meta-mode 1: Loadings show high weight on $(\lambda, \log \rho, -\tau)$ — this mode captures the "escape vs stability" axis. Points with high meta-mode 1 amplitude are deep in the interior or far outside; the boundary has low amplitude.

Interpretation of meta-mode 2: Loadings show high weight on $(\gamma, -\tau)$ — this mode captures "coherent rotation vs chaotic dephasing." Interior points with regular dynamics have high meta-mode 2; boundary points with chaotic orbits have low meta-mode 2.

5.3 Meta-Resonance Map

Region	Mean meta-resonance	Interpretation
Interior (non-escaped)	0.898	Strongly single-mode (stable dynamics)
Escaped (far exterior)	0.566	Mixed modes (transition dynamics)

Region	Mean meta-resonance	Interpretation
Boundary	0.45–0.55	Balanced between meta-modes (regime boundary)

The meta-resonance map is a cleaner regime separator than any single first-level feature.

5.4 Implementation

examples/mandelbrot_spectral_matrix_evolution.py — includes both first-level and meta-spectral computation with PNG output.

---

6. Applications

6.1 Agent Routing

First-level features: (cost, quality, uncertainty, blocker-state) per candidate action.

Meta-spectrum: identifies the dominant decision axes. If meta-mode 1 captures 80%+ of variance and loads heavily on (quality - cost), the controller reduces to a single quality-adjusted-cost score. If lead share drops, the decision landscape is genuinely multi-dimensional and requires richer reasoning.

6.2 Memory Retrieval

First-level features: (recency, semantic similarity, confidence, mode energy) per memory atom.

Meta-spectrum: reveals which retrieval signal combinations are dominant. If recency and similarity are highly correlated (high meta-mode 1 loading on both), they can be merged into a single score. If they are orthogonal (captured by different meta-modes), both dimensions must be preserved.

6.3 Proof Search

First-level features: (coherence, bridge-value, difficulty, tactic-success-rate) per proof goal.

Meta-spectrum: identifies whether the proof landscape is dominated by a single difficulty axis or requires multi-dimensional navigation. Low effective dimensionality suggests a simple priority ordering; high dimensionality suggests the need for diverse exploration strategies.

6.4 Financial Risk

First-level features: (VaR, Expected Shortfall, correlation regime, spectral gap) per asset or portfolio.

Meta-spectrum: captures the dominant risk factor interactions. In normal markets, one meta-mode (systematic risk) dominates. In crisis, effective dimensionality drops further (everything correlates) but the leading eigenvector rotates toward tail risk.

---

7. Comparison to Existing Methods

Method	What it does	Limitation
PCA	First-order eigenanalysis of raw data	No spectral pre-processing; ignores mode structure
Kernel PCA	Nonlinear first-order eigenanalysis	Single level; no meta-mode concept
Factor models	Extract latent factors from returns	Assumes linear structure; no significance test
Meta-spectrum	**Second-order eigenanalysis of spectral features**	Requires meaningful first-level features

Key advantage: the meta-spectrum inherits the optimality of its first-level spectral features. If the first level captures the system well, the second level captures how those features interact — with the same Eckart-Young guarantee.

8. Planned Experiments

8.1 Mandelbrot (Complete)

Already demonstrated. See Section 5.

8.2 Financial Regime Analysis

- First-level: spectral regime detection features (coefficient drift, spectral gap, mode energies) on daily equity returns.
- Meta-spectrum: identify how many independent regime signals exist. Compare to standard HMM regime models.

8.3 Lean Proof Graph Meta-Analysis

- First-level: per-theorem features (spectral centrality, bridge score, degree, sorry-distance).
- Meta-spectrum: identify the dominant structural patterns in the proof library. Compare meta-modes to human-identified "proof families."

8.4 Agent Decision Surface

- First-level: per-decision features from the Nous system (cost, expected gain, uncertainty, blocker count).
- Meta-spectrum: identify whether the decision landscape is effectively 1D or multi-dimensional. Use as input to controller policy design.

9. Connection to Other Papers

- **meta_mandelbrot_spectral_evolution**: First concrete validation of the meta-spectral method.
- **ml_spectral_neural_architecture**: Meta-spectrum supervises mode-gate learning; modes that form strong meta-modes should be gated together.
- **ml_spectral_cognitive_resonator**: Meta-spectrum is the controller's decision surface (L4).
- **fin_spectral_regime_detection**: Time-varying meta-spectrum provides a second-order regime-switch detector.
- **ml_spectral_intelligence**: The data spectral exponent s determines the first-level spectral structure; the meta-spectral exponent s' determines how many independent feature interactions exist.

---

10. Open Problems

1. **Higher-order meta-spectra**: Is there useful information in the third-order spectrum (spectrum of meta-spectra)? Connection to random matrix universality.

2. **Optimal first-level features**: The meta-spectrum quality depends on the choice of first-level features. Is there a principled way to select features that maximize meta-spectral informativeness?

3. **Causality**: The meta-spectrum captures correlations, not causality. Can we build a causal meta-spectrum by replacing covariance with a causal operator (e.g., Granger spectral causality applied at the feature level)?

4. **Lean formalization**: The Eckart-Young theorem is already verified in Lean. Can we extend the verification to the meta-spectral optimality theorem (Theorem 1)?