Laplacian Component Analysis (LCA)

The Role of the Regularisation Parameter as an Information - Application Bridge

May Yuan-Klitzner

January 26, 2026

1 Overview

α is not just a tuning parameter, but the mechanism that defines what we consider information in radar data. PCA assumes variance is information; LCA lets us say that *spatially coherent variance is information*, and isolated variance is often noise.

In LCA, α is the design choice that encodes how much we trust application structure such as spatial coherence in SAR relative to raw variance, and that is why LCA is iterative but more meaningful than PCA for structured radar data.

2 What LCA Adds Beyond PCA / kPCA

Goal of PCA: Maximise global variance; purely data-driven and agnostic to application structure.

Limitation: PCA treats samples as independent and ignores known relationships (spatial, temporal, physical, semantic).

Core idea of LCA: LCA augments PCA with graph-based prior information via a Laplacian regularizer. This allows the learned components to respect application-specific structure, not just statistical variance.

Mathematically, LCA balances two competing objectives:

- Statistical fidelity (variance / reconstruction)
- Geometric consistency (smoothness on a graph)

This balance is controlled by a single scalar parameter, denoted α, β, or λ, which is the central object of interest in LCA theory.

Key shift from PCA: LCA replaces variance = information with information = variance + structured smoothness.

3 The Regularisation Parameter as the "Information Dial"

3.1 Where the parameter appears

A canonical graph-regularised PCA / LCA objective is:

$$\min_{U,Q} \|X - UQ^T\|_F^2 \; + \; \alpha \, \mathrm{Tr}(Q^T L Q), \quad \text{s.t. } Q^T Q = I$$

where:

- $X \in \mathbb{R}^{d\times n}$ is the data matrix,
- $L = D - W$ is the graph Laplacian,
- W encodes similarity or adjacency between samples,
- Q are the low-dimensional representations,
- $\alpha \geq 0$ is the regularisation parameter.

This formulation balances two competing objectives:

- **Term 1: PCA term:**
$$\|X - UQ^T\|_F^2$$
⇒ preserves global variance / reconstruction accuracy
⇒ Wants: Capture all the variance
⇒ Cares about: Statistical patterns
- **Term 2: Laplacian Term**
$$\mathrm{Tr}(Q^T L Q)$$
⇒ enforces smoothness with respect to a graph W (similarity matrix)
⇒ Wants: Connected points to have similar representations
⇒ Cares about: Geometric structure
- **α: The balance knob**
⇒ Controls: How much we weight structure vs. variance
⇒ Range: 0 to ∞
⇒ Effect: Defines what "information" means for YOUR application

3.2 Geometric meaning of the Laplacian term

For a single embedding co-ordinate $q \in \mathbb{R}^n$:

$$q^T L q = \frac{1}{2}\sum_{i,j} W_{ij}(q_i - q_j)^2$$

Interpretation:

- For every pair of points (i, j), if they're connected in the graph (W_{ij} is large), penalize them for having different representations ($q_i \neq q_j$).
- Therefore, strongly connected points (W_{ij} large) are penalised for having very different representations.
- Minimization encourages low-frequency (smooth) signals on the graph.
- High-frequency variations correspond to noise or sharp, local inconsistencies.

Thus, the Laplacian term acts as a smoothness enforcer on the data manifold.

3.3 Trade-off interpretation of α

The regularisation parameter explicitly controls what you consider "information":

- $\alpha = 0$**: Pure PCA**
  $\Rightarrow$ information = variance only
  $\Rightarrow$ Use when data points are truly independent and there is no spatial structure
  $\Rightarrow$ Example: Analyzing completely independent radar measurements

- $\alpha \to \infty$**: Pure Laplacian embedding**
  $\Rightarrow$ information = graph geometry only
  $\Rightarrow$ Use when variance is dominated by noise and graph structure is very reliable
  $\Rightarrow$ Example: Extremely noisy environment (low SNR)

- **Intermediate** α**:**
  $\Rightarrow$ Components explain variance while remaining smooth on the graph
  $\Rightarrow$ Example: Need both target discrimination AND clutter suppression

Many formulations normalise the objective and reparameterise using $\beta \in [0, 1]$, making this trade-off explicit:

- $\beta = 0$: PCA
- $\beta = 1$: Laplacian eigenmaps
- β: fraction of weight given to application-specific structure

4 Why LCA Is Iterative

Unlike PCA, which admits a direct closed-form solution via eigendecomposition of the covariance matrix, LCA couples reconstruction and graph smoothness. This coupling typically leads to alternating or iterative optimisation procedures rather than a single closed-form solution.

5 Application-Specific Meaning of α

Choosing α is equivalent to deciding what is "signal" and what is "noise" in YOUR application.

Example: Radar Remote Sensing

- **Small** α ($\alpha \approx 0.01$)
  What it means: "Variance is information, smoothness is less important"
  Use for:
  - Target detection
  - Finding anomalies

- Discriminating different target types

Assumes:
- Targets create variance (stand out from background)
- Less concern about spatial coherence
- Want to preserve differences

Example scenario: Detecting vehicles in an open field
- Vehicles are different from grass (high variance = good!)
- Don't need strong spatial smoothing
- Small α preserves these discriminative features

- **Large α ($\alpha \approx 10$)**
  What it means: "Spatial structure is information, variance might be noise"
  Use for:
  - Clutter suppression
  - Denoising speckle
  - Scene understanding
  - Terrain classification

  Assumes:
  - True signal is spatially coherent
  - Isolated variations are likely noise
  - Neighbors should be similar

  Example scenario: Mapping forest canopy structure
  - Trees create spatially coherent regions
  - Speckle noise is random (high variance = bad!)
  - Large α enforces smoothness, suppresses speckle

- **Medium α ($\alpha \approx 1$)**
  What it means: "Balance variance and structure"
  Use for:
  - General-purpose feature extraction
  - When both global and local matter
  - Classification tasks

  Assumes:
  - Some variance is signal, some is noise
  - Some structure is real, some is overfitting
  - Need balanced representation

  Example scenario: Urban scene analysis
  - Buildings have sharp edges (preserve variance)
  - But also spatial coherence within structures (enforce smoothness)
  - Medium α balances both needs

---

6 Summary:

What is α?

- **Mathematically:**
  - A regularisation parameter
  - Controls trade-off between variance and smoothness
  - Appears in front of Laplacian term
- **Conceptually:**
  - An information encoder
  - Defines what you consider signal vs. noise
  - Injects domain knowledge into algorithm
- **Practically:**
  - An application tuning knob
  - Small $\rightarrow$ emphasize discrimination
  - Large $\rightarrow$ emphasize coherence
  - Medium $\rightarrow$ balance both