# Shape-based data representations for experimental spectra

## Applications to material retrosynthesis and phase mapping


Kiran Vaddi, Huat T Chiang, Karen Li, and Lilo D Pozzo

kiranvad@uw.edu    kiranvad.github.io


Inculcation of shape based statistical models improve efficacy of machine learning on experimental data.

Spectral data are ubiquitous in the experimental material design and discovery for structure and performance characterizations. Examples include UV-Vis spectroscopy for plasmonic resonance characterization of gold nanoparticles and X-ray scattering of crystalline and lipid phase behavior. Performing material retrosynthesis of target structures or high-throughput phase mapping using a material acceleration platform would require the robotic agent to manipulate and compare spectral data such that the shape of the spectra and deformations are represented appropriately.

### Function space representations

- Spectral data are functions mapping a stimuli to response rather than high dimensional vectors
- Simple measures of distance or mean of functions needs careful handling of information on x and y axis
- Manifold based statistics and learning methods provide ways to encode and process shape information : slope transitions, peak shifts and broadening encoded in function

## Amplitude-Phase distance

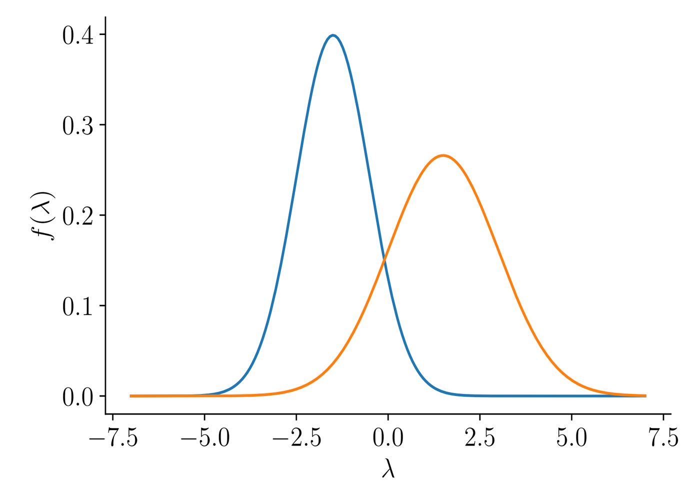
DOI 10.1007/978-1-4939-4020-2

A *metric-based representation* to encode shape information via distance to be used in defining reward functions and clustering

- Define SRSF of a function using:

Encapsulate shapes via derivative

$$q(\lambda) = \dot{f}(\lambda)/\sqrt{\left|\dot{f}(\lambda)\right|}$$

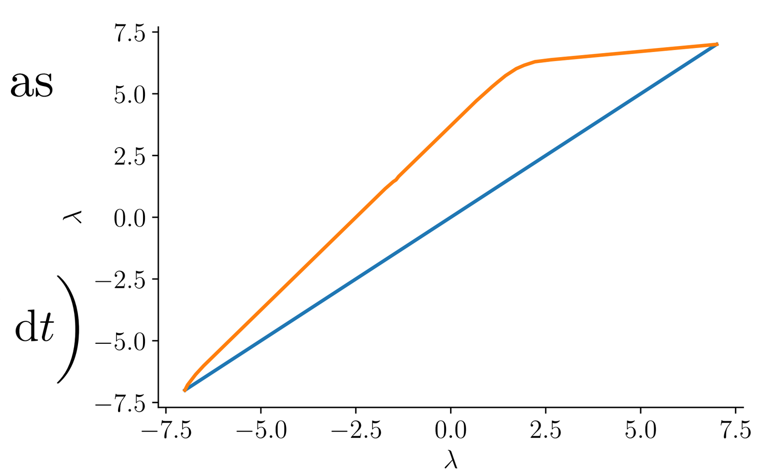



Functions with phase and amplitude variation

- Phase information in warping function:

Align features such as peaks and valleys

$$d_p = \cos^{-1}\left(\int_0^1 \sqrt{\dot{\gamma}(t)}\ \mathrm{d}t\right)$$

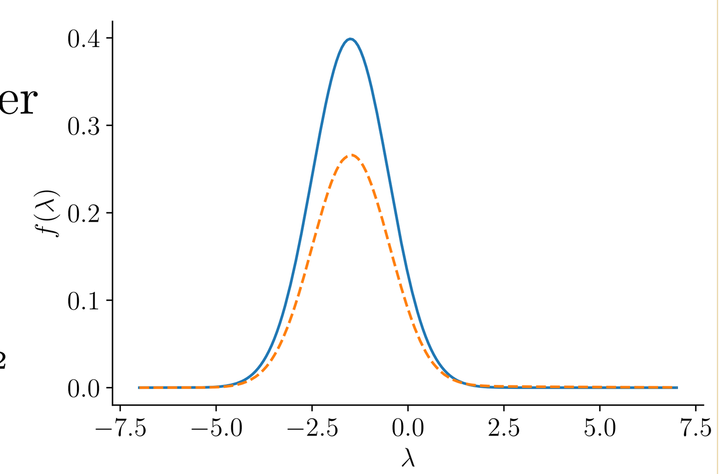


Optimal alignment along x-axis

- Shape information in aligned functions :

Amplitude variation after alignment is the shape mismatch

$$d_a = \inf_{\gamma} \|q_1 - (q_2, \gamma)\|_{\mathbb{L}^2}$$



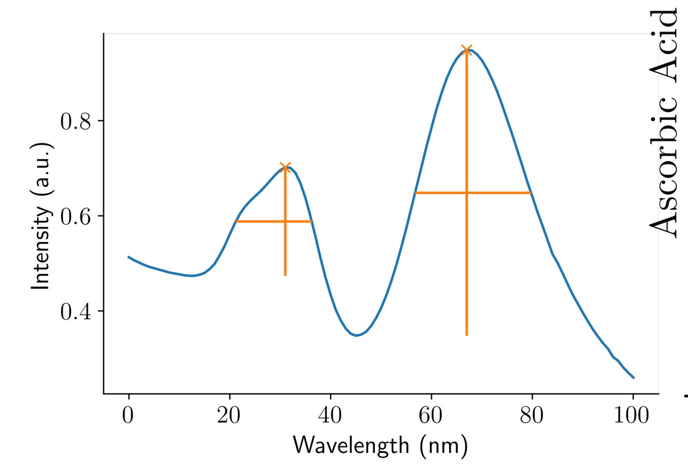

Aligned functions

## Gold-nanorods retrosynthesis

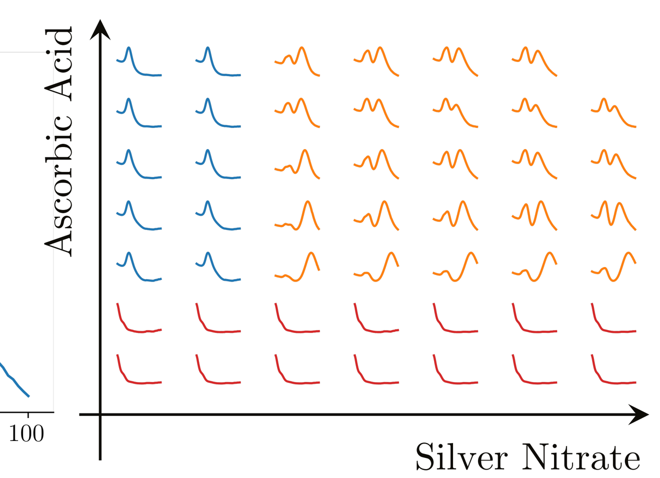
DOI 10.1039/D2DD00025C

Goal : Autonomously retrosynthesize gold nanorods by matching shape of plasmonic resonance signal in UV-Vis spectroscopy

- Two-dimensional synthesis space :

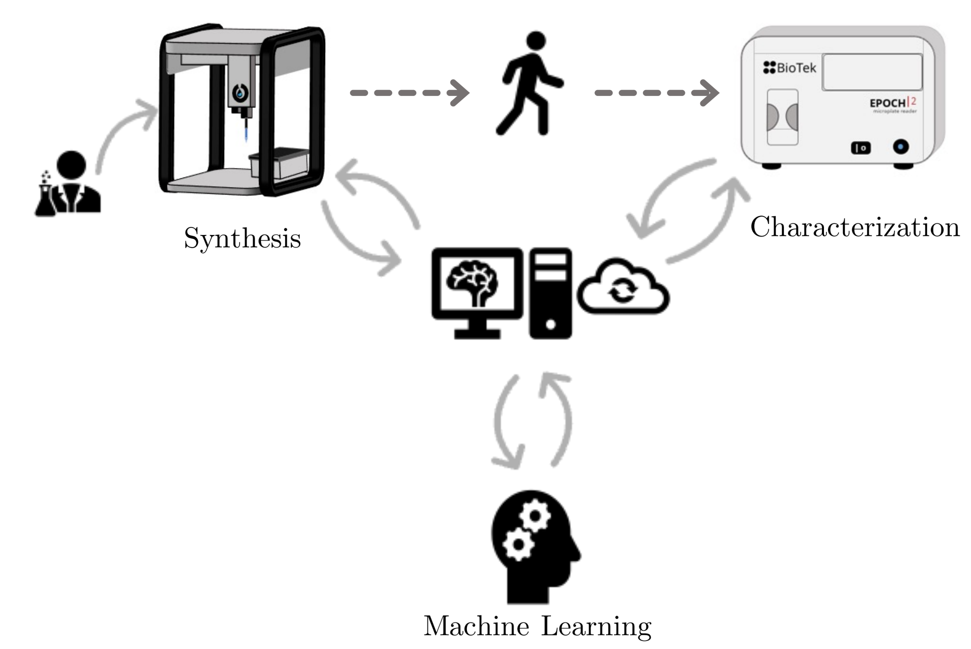



Target spectrum

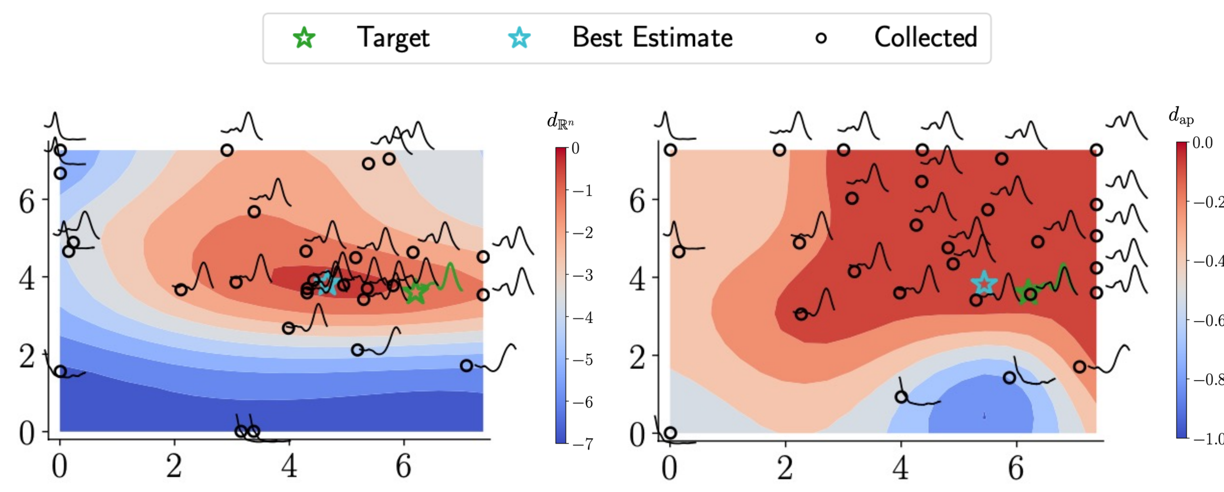



Ground truth state diagram

- Closed-loop experimentation :




- Amplitude-Phase learns state diagram:



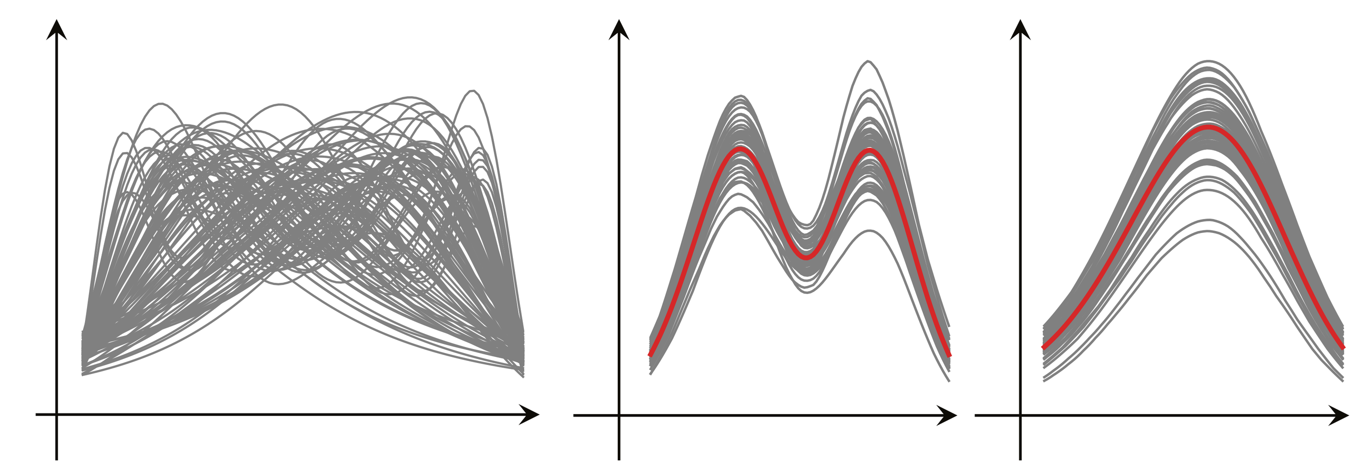

Euclidean    Amplitude-Phase

## Phase mapping of block-copolymers

arXiv:2011.12397

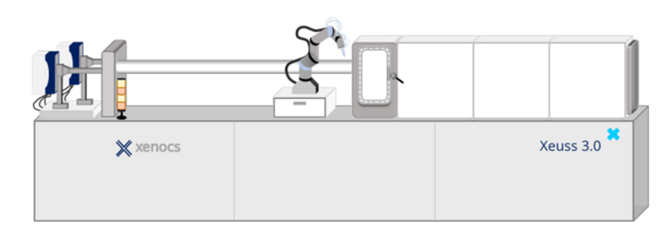
Goal : Summarize possible phases formed in self-assembled block-copolymers by global template learning based on the shape of a SAXS curve

- Peak shift and widths are shape independent :

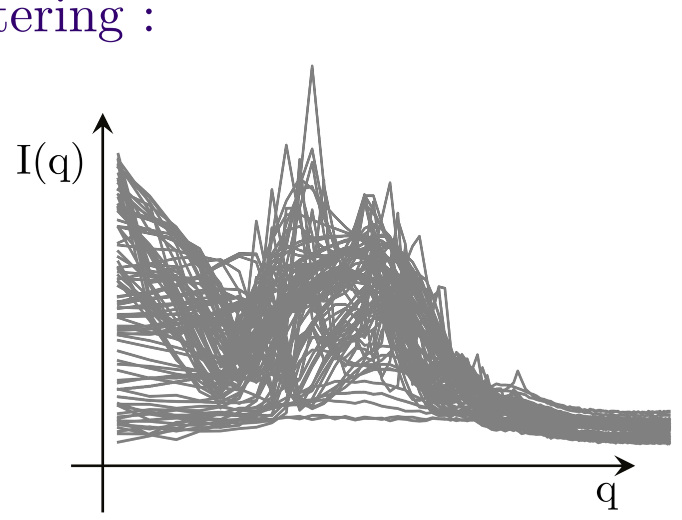


Functional data    Clusters in functional representation

- Small Angle X-ray Scattering :



High-throughput SAXS




SAXS data in functional form

- Learned templates are different phases :