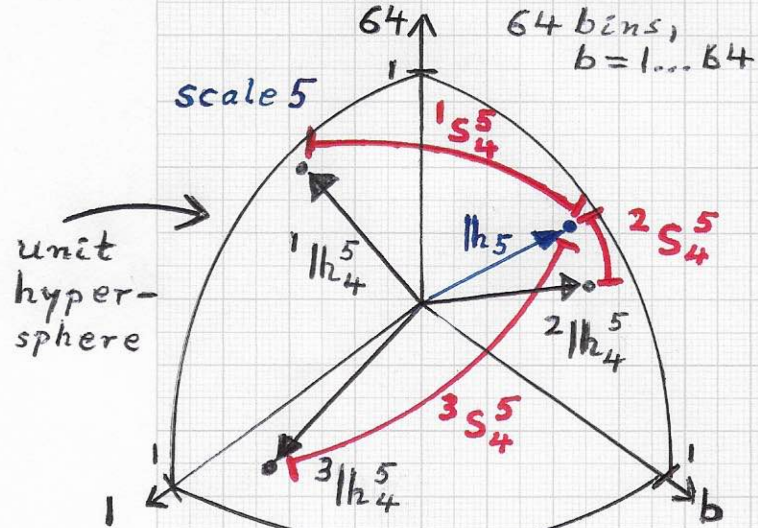
OBJECT AND MATERIAL EIGENFUNCTIONS - Cont'd.

- Laplacian eigenfunctions and neural networks:... • Using one of the two norms mentioned at the bottom of the previous page (or some other appropriate norm), the MAX norm $\|\cdot\|_{max}$ emphasizes the one scale of the to-be-classified segment that is maximally similar to the same scale of a specific class and class segment, while the L_2 norm $\|\cdot\|_2^2$ considers and sums up squared similarities between all scales of the to-be-classified segment and all scales of a specific class and class segment.

Examples. The indices cl (class), sg (segment) and sc (scale) and the number of bins of histograms ("histogram dimension") can be confusing when describing the method using the general notation ${}^{sg}\mathbb{h}_{cl}^{sc}$ for a classified histogram in the database and $\mathbb{h}_{sc}$ for a scale-sc histogram of the to-be-classified segment



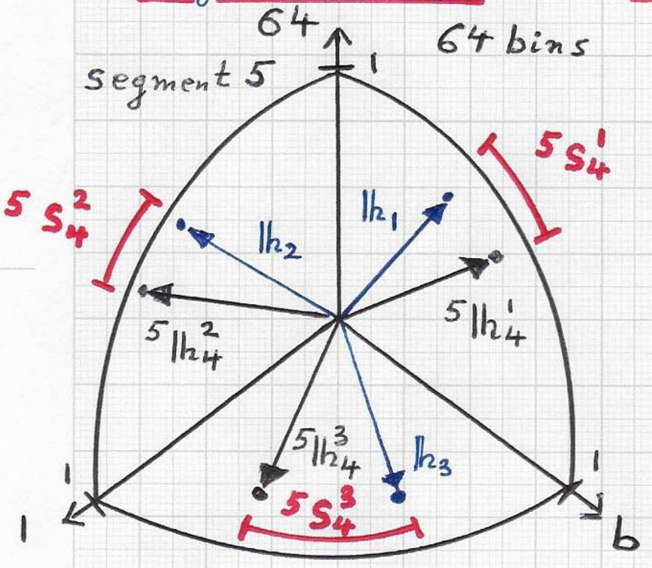
class 4, segments 1,2,3 — varying segments.

The figure (left) shows an example: Scale-5 similarity values for 3 segments (1,2,3) representing class 4. Three COS similarity values result: ${}^1S_4^5$, ${}^2S_4^5$ and ${}^3S_4^5$.

...

OBJECT AND MATERIAL EIGENFUNCTIONS – Cont'd.

• **Laplacian eigenfunctions and neural networks:...** In this example, the similarity value ${}^{2}s_4^{5}$ is maximal, since the inner product $\langle \mathbb{h}_5, {}^{2}\mathbb{h}_4^{5} \rangle$ is maximal. Thus, the scenario sketched in the figure on the previous page keeps class (4) and scale (5) fixed and varies the segments (1,2,3). The next example keeps class (4) and segment (5) fixed and varies the scales (1,2,3).



class 4, scales 1, 2, 3 – Varying scales.

The figure (left) shows an example where scale-1, scale-2 and scale-3 similarity values for one segment (5), being one of a representative set of segments, are calculated for class 4. The three resulting COS similarity values are: ${}^{5}s_4^{1}$, ${}^{5}s_4^{2}$ and ${}^{5}s_4^{3}$.

This second scenario is rather relevant for our material data classification application:

THE DIFFERENT SCALES OF THE TO-BE-CLASSIFIED SEGMENT HAVE DIFFERENT DEGREES OF SIMILARITY FOR THE SAME SCALES OF A CLASSIFIED SEGMENT (of a specific class).

OBJECT AND MATERIAL EIGENFUNCTIONS - Cont'd.

- Laplacian eigenfunctions and neural networks:... For this last special scenario, one obtains ${}^5\mathbb{S}_4 = ({}^5s_4^1, {}^5s_4^2, {}^5s_4^3)$. Thus, the two suggested norms for ${}^5\mathbb{S}_4$ have the values

  i) $\|{}^5\mathbb{S}_4\|_{max} = \max\{{}^5s_4^1, {}^5s_4^2, {}^5s_4^3\}$ and

  ii) $\|{}^5\mathbb{S}_4\|_2^2 = ({}^5s_4^1)^2 + ({}^5s_4^2)^2 + ({}^5s_4^3)^2$.

  - **Note.** The value of $\|{}^5\mathbb{S}_4\|_{max}$ is in the interval $[0,1]$. The value of $\|{}^5\mathbb{S}_4\|_2^2$ is in the interval $[0,3]$ and should be normalized to the interval $[0,1]$ by dividing it by the number of scales used for its computation, i.e., three in this case.

- Based on the vector norms i) and ii), one can define an overall degree of similarity between the to-be-classified segment and each class in the database. For this overall similarity one must consider all scales of all the segments stored for a class — to compute the value of overall similarity with this class, called S_{cl}, $cl=1\ldots C$. One obtains these two definitions of S_{cl}:

  i) $$\underset{cl=1\ldots C}{S_{cl}} = \max_{sg=1\ldots k_{cl}} \left\{ \max_{sc=1\ldots H} \left\{ {}^1s_{cl}^1, \ldots, {}^1s_{cl}^H \right\}, \ldots, \max_{sc=1\ldots H} \left\{ {}^{k_{cl}}s_{cl}^1, \ldots, {}^{k_{cl}}s_{cl}^H \right\} \right\}.$$

  "Overall similarity with class cl = maximum of all maxima of all similarities with all k_{cl} segments and all H scales stored for class cl."

...

OBJECT AND MATERIAL EIGENFUNCTIONS – Cont'd.

• Laplacian eigenfunctions and neural networks:... In other words, the value of norm i) is based only on the largest observed similarity with all class segments across all scales.

ii) $$\underline{S_{cl}}_{cl=1\ldots C} = \max_{sg=1\ldots k_{cl}} \left\{ \frac{1}{H}\sum_{sc=1}^{H} \left({}^{1}s_{cl}^{sc}\right)^2, \ldots, \frac{1}{H}\sum_{sc=1}^{H} \left({}^{k_{cl}}s_{cl}^{sc}\right)^2 \right\}.$$

"Overall similarity with class cl = maximum of all similarities with all k_{cl} segments, via combining all H scale similarities, for class cl." In other words, the value of norm ii) is based on an 'average' of H similarity values with all segments representing a specific class.

• Once the decision is made concerning the most appropriate definition of a norm for overall similarity computation, one can calculate S_{cl} values for C classes for the to-be-classified segment. One obtains an overall "similarity vector"

$$\mathbb{S} = (S_1, \ldots, S_C).$$

• Crucially important for the performance of all subsequent data analysis and classifcation steps is the relationship between $\mathbb{S}$ and the probability of belonging to a certain class cl.

...

Stratovan

OBJECT AND MATERIAL EIGENFUNCTIONS – Cont'd.

• Laplacian eigenfunctions and neural networks:...

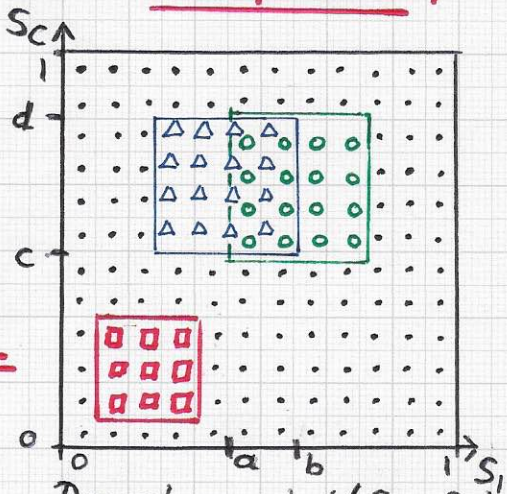
• The mapping of S_{cl}-values to class-match probability values is extremely challenging. This mapping cannot be defined from a purely theoretical perspective; THE MAPPING MUST BE ESTABLISHED VIA THE USE OF LARGE "REAL-WORLD DATASETS." Ultimately, we must define a "near-optimal mapping"

$$(S_1, \ldots, S_C) \mapsto (p_0, p_1, \ldots, p_C).$$

The probabilities p_{cl}, $cl = 0 \ldots C$, define the likelihood that a to-be-classified segment belongs to one of the $(C+1)$ classes; class 0, as discussed before, is the class "different from all classes $1 \ldots C$. One must ensure that

$$0 \leq p_{cl} \leq 1,\ cl = 0 \ldots C, \text{ and } \sum_{cl=0}^{C} p_{cl} = 1.$$

Here: $(S_1, S_C) \in [a,b] \times [c,d] \Rightarrow p_2 = 0.5,\ p_3 = 0.5.$



Densely sampled $(S_1, \ldots, S_C)$ domain. Four classes are shown: 0=•, 1=□, 2=△, 3=○. Classes 2 and 3 overlap in the $\mathbb{S}$ domain, leading to reduced match probabilities where they overlap.

• Once a database exists with a finite number of materials data, one would "ideally" have available an INFINITE amount of additional classified segments; compute the $\mathbb{S}$ vectors for these segments; and establish an INFINITELY DENSE SAMPLING of corresponding tuples $\mathbb{S}$ and the known class labels cl, $cl \in \{0, \ldots, C\}$. The left figure shows an example.

...