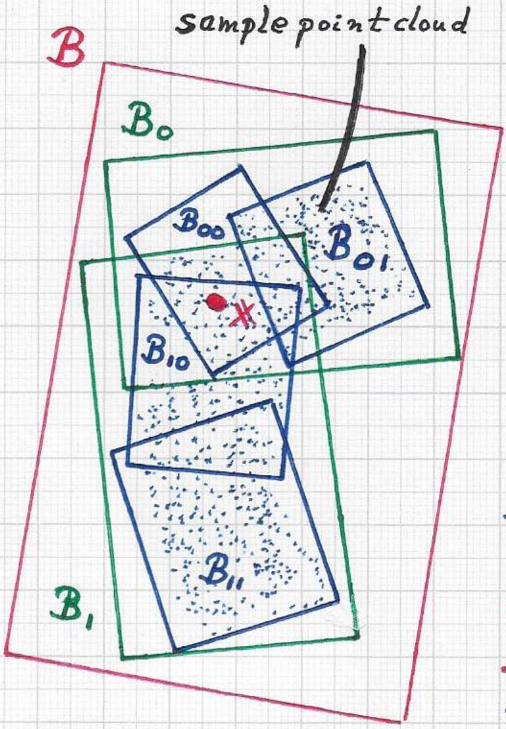
OBJECT AND MATERIAL EIGENFUNCTIONS – Cont'd.

- Laplacian eigenfunctions and neural networks:...



We can view the construction the described bounding box tree structure as a global optimization problem: The tree structure must support the overall optimization objective of calculating optimal **P**-values, i.e., optimal probability values for class-membership. The left figure is a sketch of the box geometry of a tree with root node B having the two children B_0 (with the two children B_{00} and B_{01}) and B_1 (with the two children B_{10} and B_{11}). The described tree construction method is inherently deterministic, when computing the geometrical characterists of boxes (when performing repeated binary box splitting). The geometry of each box is defined, for example, by its center and its orientation and scale, given by the D box basis vectors (not normalized). Thus, the **P**-value computed for a to-be-classified point $*$ depends on these geometrical characteristics of all boxes in the box tree. Therefore, one can understand the computation of the geometry of all boxes as a global optimization problem.

...

OBJECT AND MATERIAL EIGENFUNCTIONS - Cont'd.

- Laplacian eigenfunctions and neural networks: ... The presented method for deterministic box tree construction is a "greedy" method – and does not pursue the goal of attempting to generate a globally near-optimal box tree (geometry), leading to near-optimal **P**-value computation for a specified number of boxes to be produced. Consequently, the "quality" of a box tree is defined by the "quality" of **P**-values that impact the overall classification performance metric. One can consider the use and adaptation of various global, combinatorial optimization approaches for the near-optimal computation of the geometry of a fixed number of boxes. For example, one can consider **GENETIC ALGORITHMS** – a specific class of evolutionary algorithms – for this purpose. The **P**-value computed for a to-be-classified point ∗ represents the point's class-membership probability (in a general multi-class classification setting). Therefore, the **P**-values computed impact the numbers of TNs, TPs, FNs, FPs and MCs (mis-classifications) – that determine the value of the function used to calculate the classification system's overall performance, as discussed before. The goal is to maximize the performance function via global optimization of the box tree.

...

OBJECT AND MATERIAL EIGENFUNCTIONS – Cont'd.

• Laplacian eigenfunctions and neural networks: ... In the following, basic concepts used in the design of genetic algorithms (GAs) are reviewed at a high level. The description of these concepts is driven by the goal of optimizing probability (P-value) computations via a hierarchical, binary bounding box data structure. The essential concepts are:

- **Population.** When adapting GA concepts to our setting, a population consists of a set of individual members, i.e., a set of box trees.
- **Fitness.** Each population member has an associated fitness value. In our scenario, the fitness of a box tree is the classification system's performance value resulting when using this boxtree for classification.
- **Offspring.** Members of a population (usually member pairs) can produce new, next-generation members by "combining" their characteristics. The goal is to improve overall fitness.
- **Selection.** The fitness values of individual members are used to establish an order for the members, with the fittest members at the top of this order. When producing offspring, members with higher fitness values are prioritized for producing offspring. ...

■ OBJECT AND MATERIAL EIGENFUNCTIONS – Cont'd.

• Laplacian eigenfunctions and neural networks: ...

• **Elimination.** Generally, the population size should be constant, i.e., when an offspring production step introduces K new members, then the GA will have to eliminate K old members.

• **Combination / Recombination.** When using two selected members for offspring production, the characteristics of the selected members are combined to establish the offspring's characteristics. In the biological context, this step can be viewed or implemented as the cross-over recombination process – where the term crossover refers to the selection and new combination of parts of the members' chromosomes.

• **Mutation.** Considering our driving application, i.e., producing offspring box trees, mutation can be understood as the random change of one (or multiple) characteristic(s) of a new offspring member. Again, this term refers to a biological phenomenon, the random change(s) in an offspring's genetic make-up.

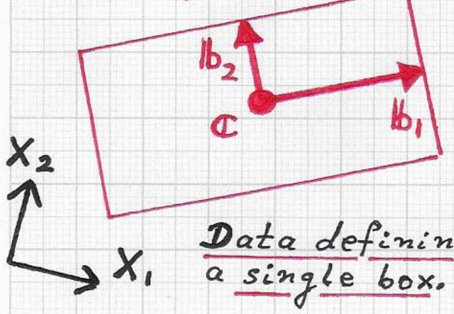
• **Characteristics.** In our case, the characteristics of a member, i.e., of an individual box tree, are the center points and the D basis vectors (mutually orthogonal, not normalized) of the boxes.

• • •

OBJECT AND MATERIAL EIGENFUNCTIONS–Cont'd.

- Laplacian eigenfunctions and neural networks:...

D = 2



Data defining a single box.

- **Characteristics.** Assuming that a box tree is viewed as a tree obtained after performing repeated binary box splitting L times for a given initial box $\mathbb{B}$, the total number of boxes in the tree (tree nodes) is $1+2+4+\ldots+2^L = 2^{L+1}-1$. Therefore, considering boxes in D-dimensional space, the number of center points is $2^{L+1}-1$ and the number of basis vectors is $D\cdot(2^{L+1}-1)$, completely defining the geometry of one box tree. The top figure shows the data defining the geometry of one box.
- **Initial population/Final population.** Given the initial sample point cloud  representing the class being considered, one must establish an initial box tree population consisting of T trees. For example, one can consider the design of a method that combines deterministic and random step to create tree geometry. The initial population — generation 0 — is subsequently used to perform selection and offspring-generating operations, thereby producing generations 1, 2, 3,... The GA terminates when a population has a member leading to the required classification performance.

•••