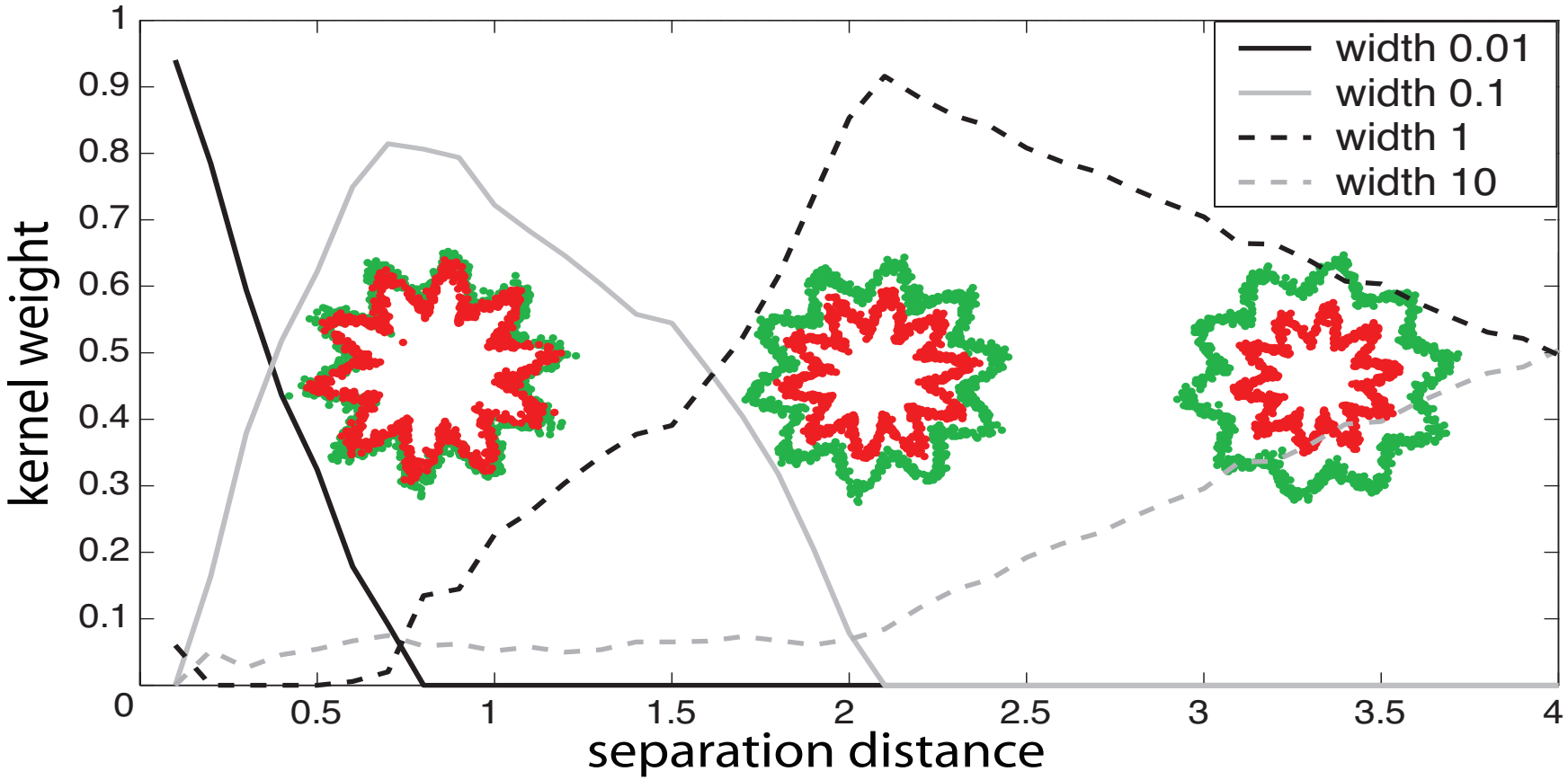
A General and Efficient Multiple Kernel Learning Algorithm

Sören Sonnenburg, Gunnar Rätsch, Christin Schäfer



$$f(\mathbf{x}) = \text{sign}\left(\sum_{i=1}^{N} \boldsymbol{\alpha_i} y_i \mathbf{k}(\mathbf{x_i}, \mathbf{x}) + \boldsymbol{b}\right) \quad \text{with} \quad \mathbf{k}(\mathbf{x_i}, \mathbf{x}) = \sum_{k=1}^{K} \boldsymbol{\beta_k} \mathbf{k}_k(\mathbf{x}_i, \mathbf{x})$$

Trick: Semi-Infinite Linear Programming

Why?

- Tool to interpret SVM Classifier
- Combine heterogeneous Data
- Automated Model Selection, Knowledge Discovery

Our contributions:

1. Efficient implementation
   - applicable to datasets of size up to 1 million with few kernels and to datasets of size around 1000 with 400 kernels.
2. General formulation
   - for very different loss functions
   - on poster: classification, regression, one-class classification.