Multidimensional Data Visualization

Dimensionality of Embedding Space

Multidimensional scaling (MDS) is a difficult global optimization problem

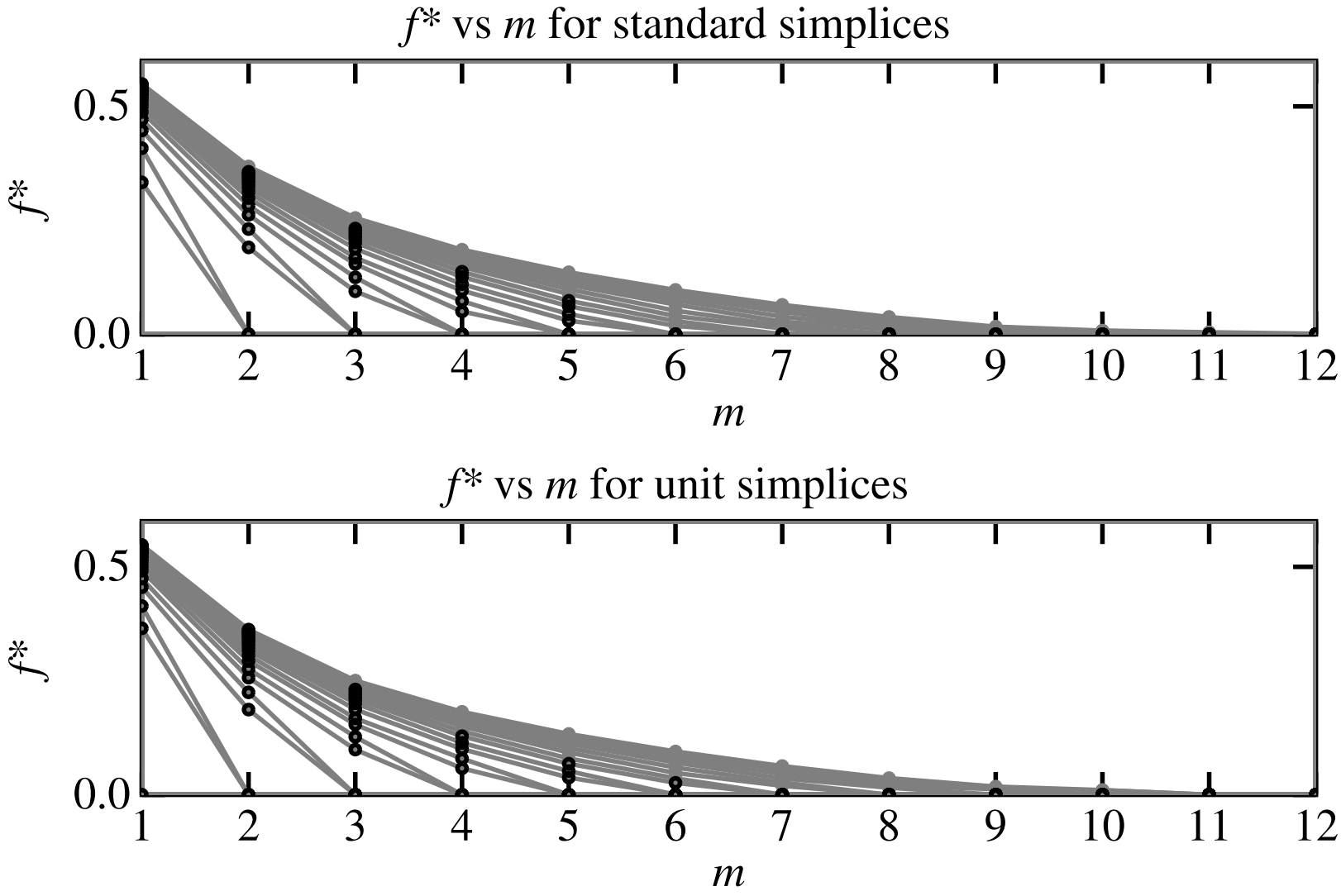
- The points representing objects should be found whose inter-point distances fit the given dissimilarities.
- The problem is reduced to minimization of a fitness criterion, e.g. so called *Stress* function

$$S(\mathbf{x}) = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \left(d\left(\mathbf{x}_i, \mathbf{x}_j\right) - \delta_{ij} \right)^2 .$$

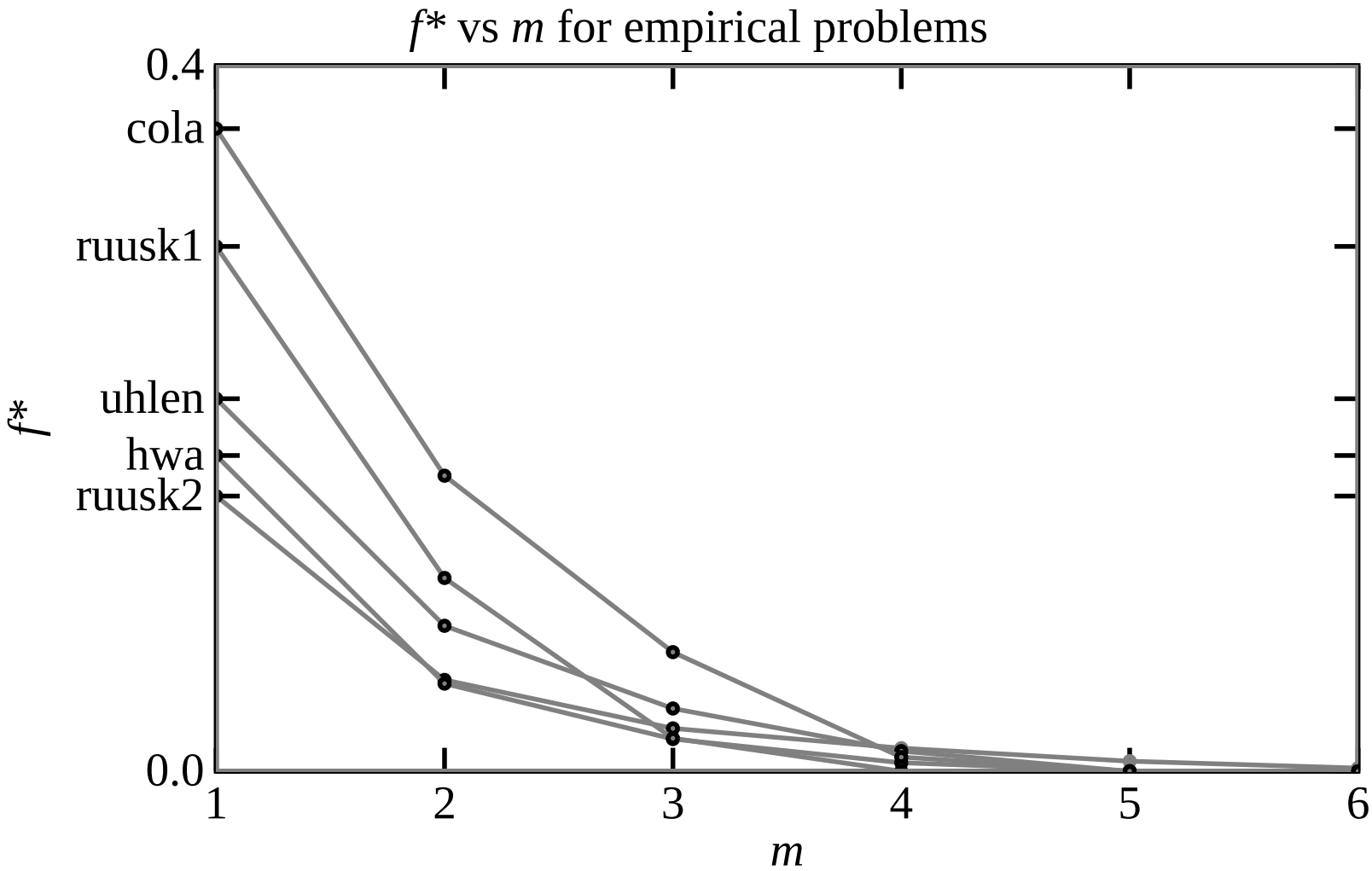
- Although *Stress* function seems rather simple, it normally has many local minima.
- The problem is high dimensional: $\mathbf{x} \in \mathbb{R}^N$ and the number of variables is equal to $N = n \times m$.
- Non-differentiability normally cannot be ignored. Minkowski distances:

$$d_r(\mathbf{x}_i, \mathbf{x}_j) = \left(\sum_{k=1}^{m} \left| x_{ik} - x_{jk} \right|^r \right)^{1/r} .$$

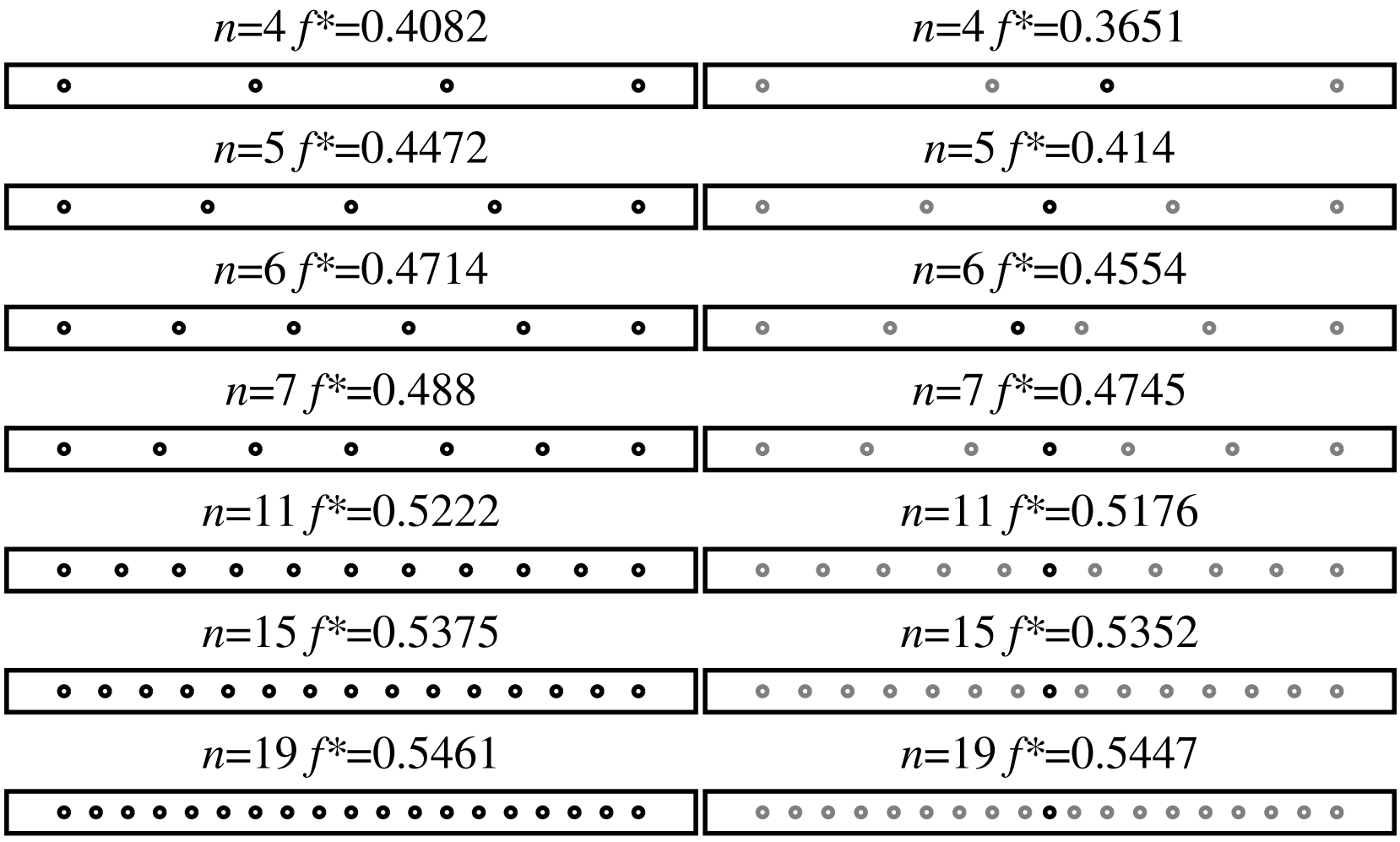
On dimensionality of embedding space



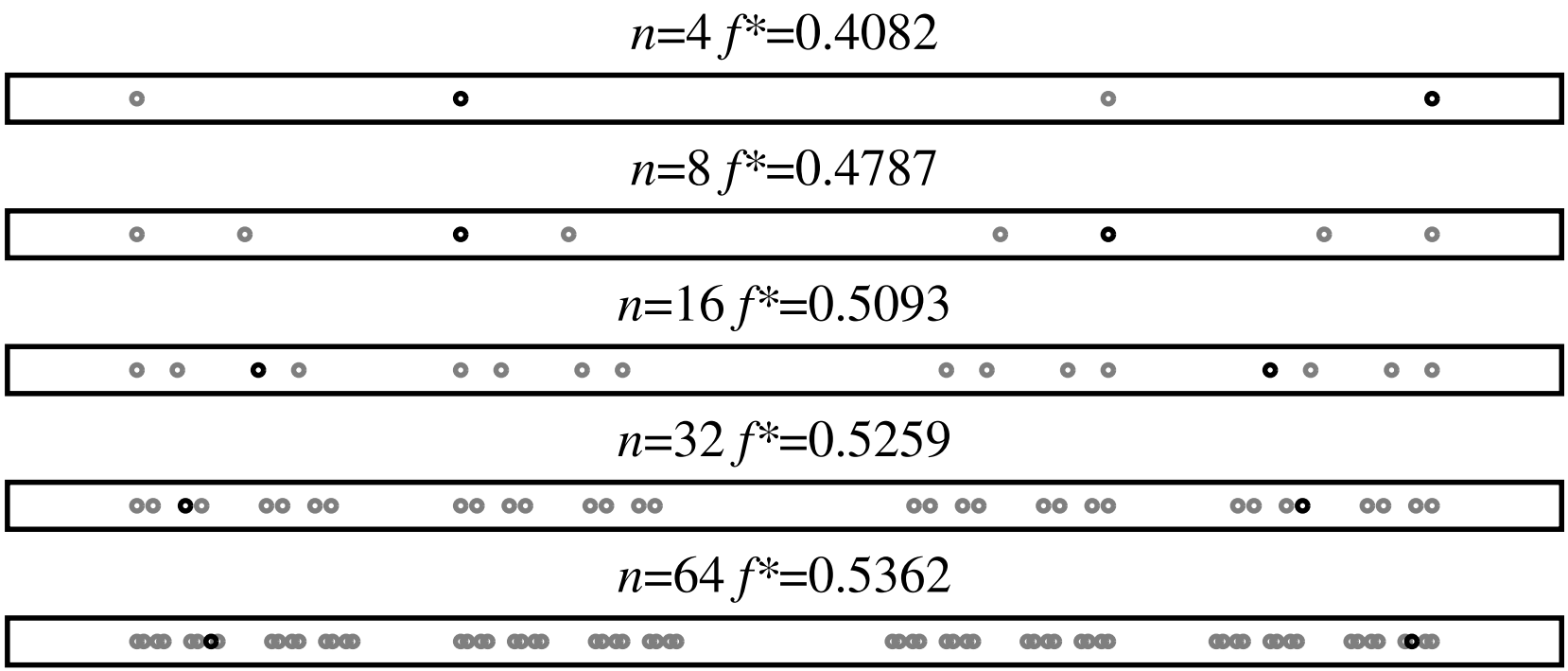
On dimensionality of embedding space



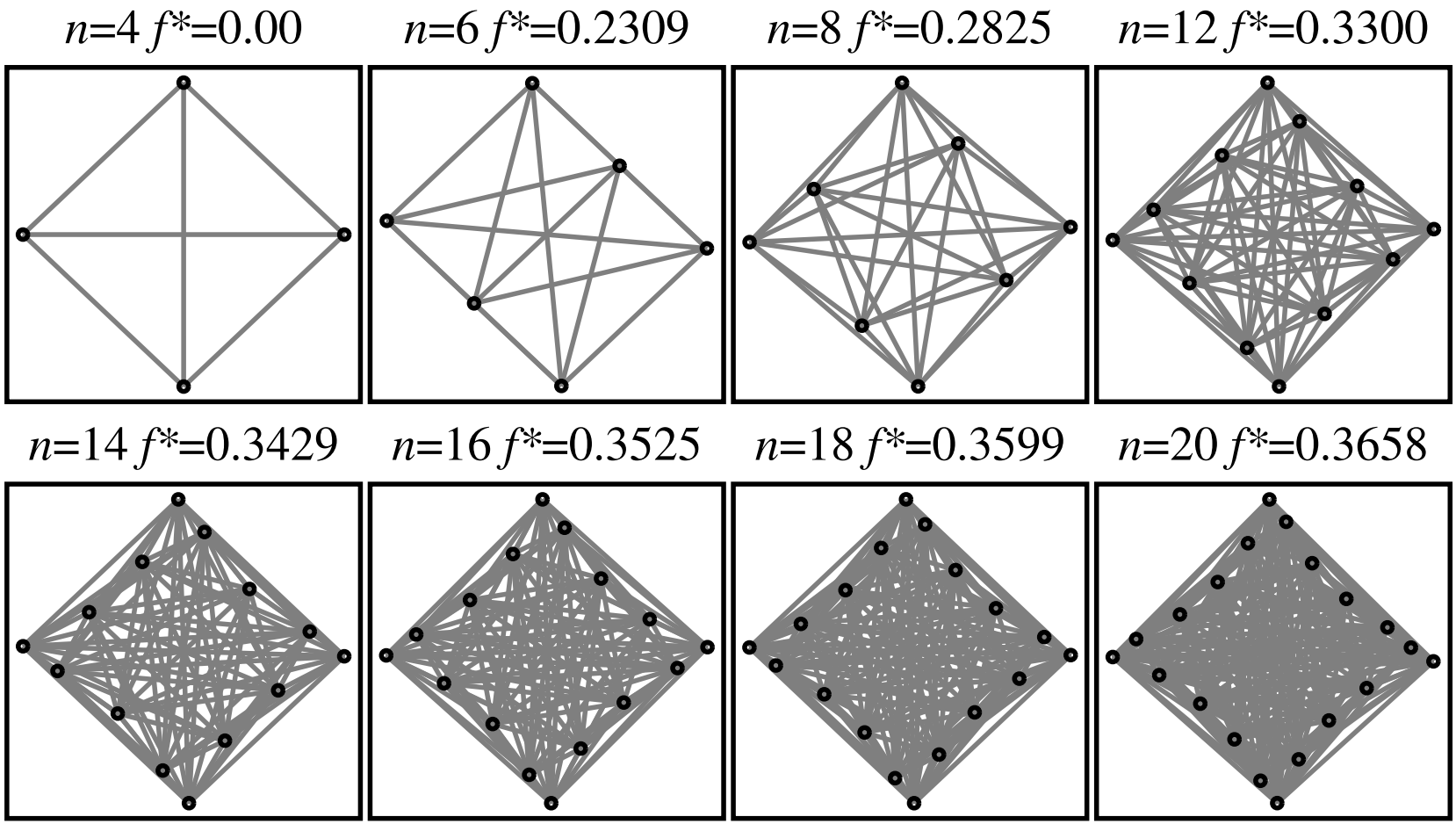
Images of simplices, $m = 1$



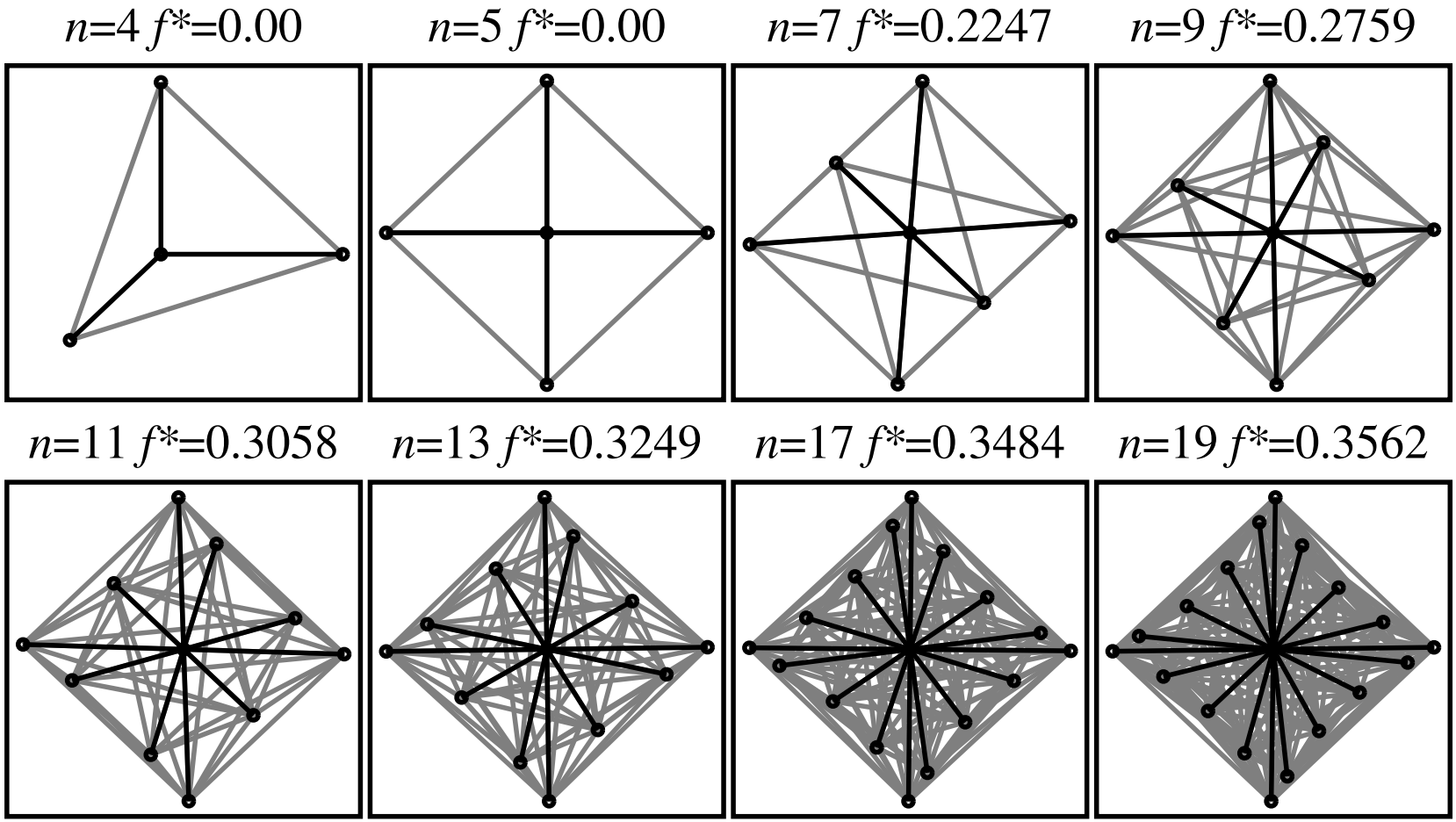
Images of cubes, $m = 1$



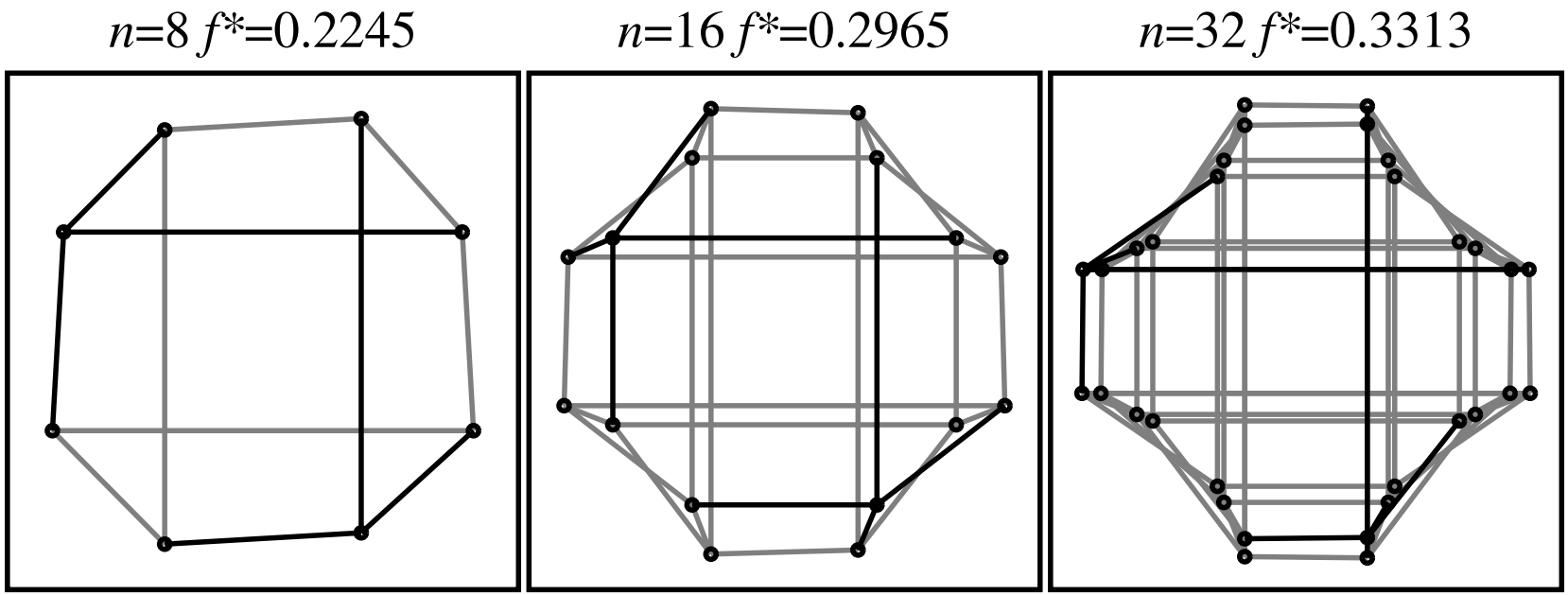
Images of standard simplices, $m = 2$



Images of unit simplices, $m = 2$

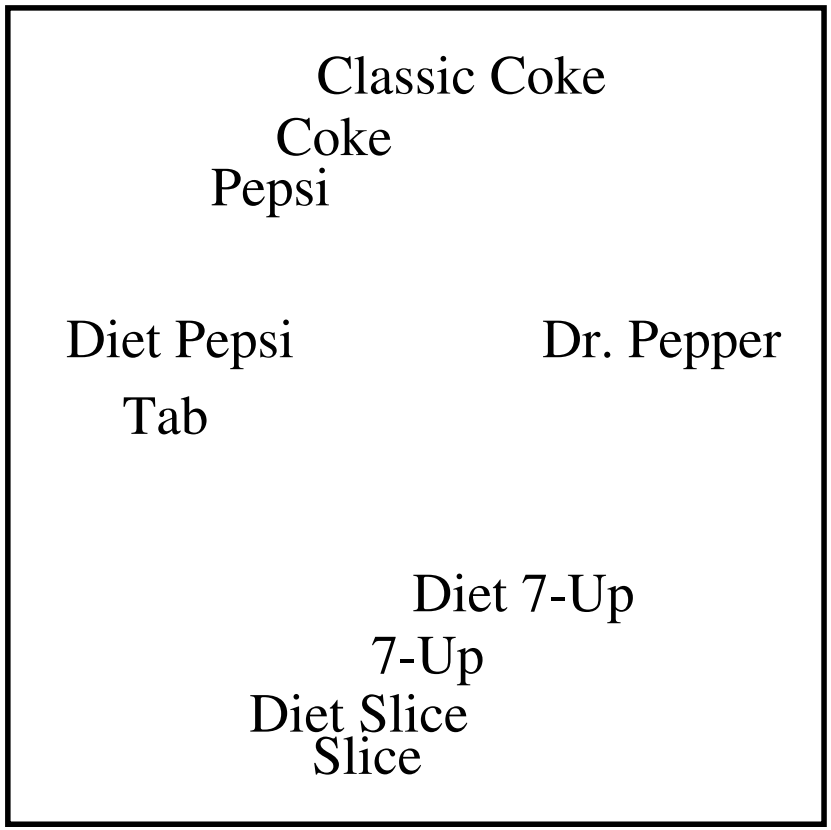


Images of cubes, $m = 2$

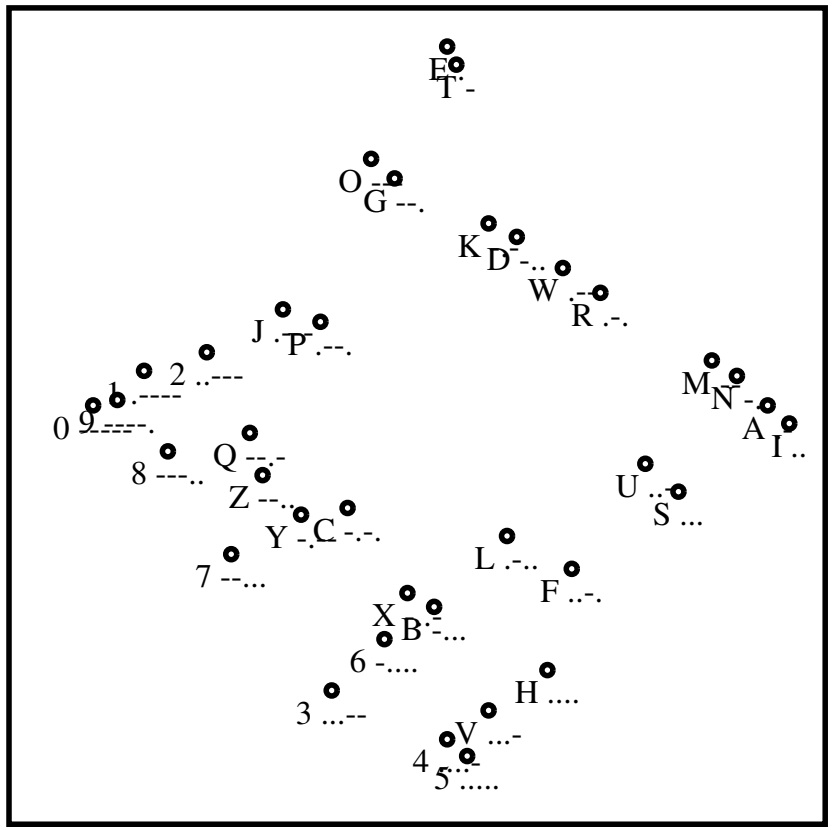


Images of empirical data, $m = 2$

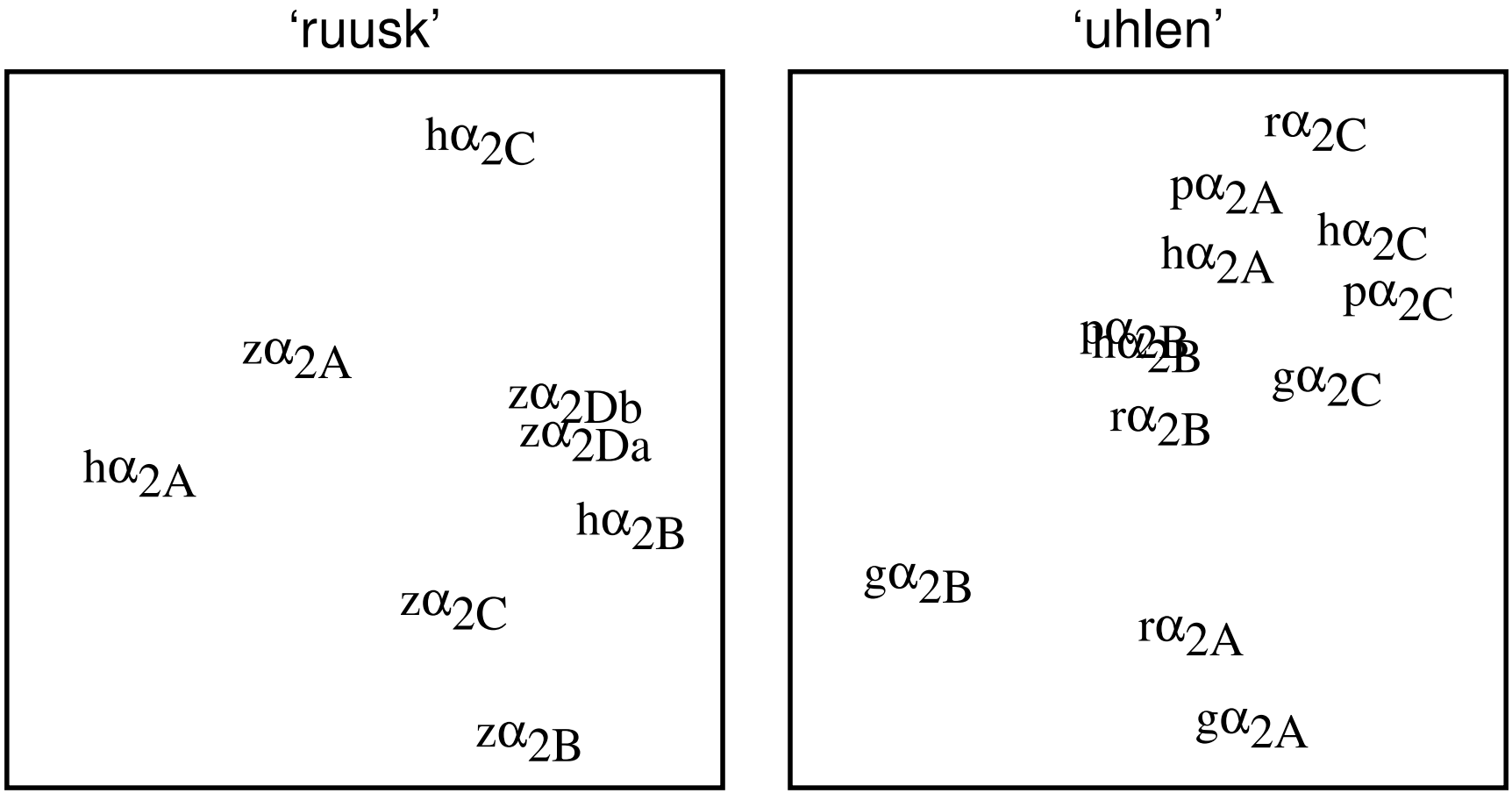
‘cola’



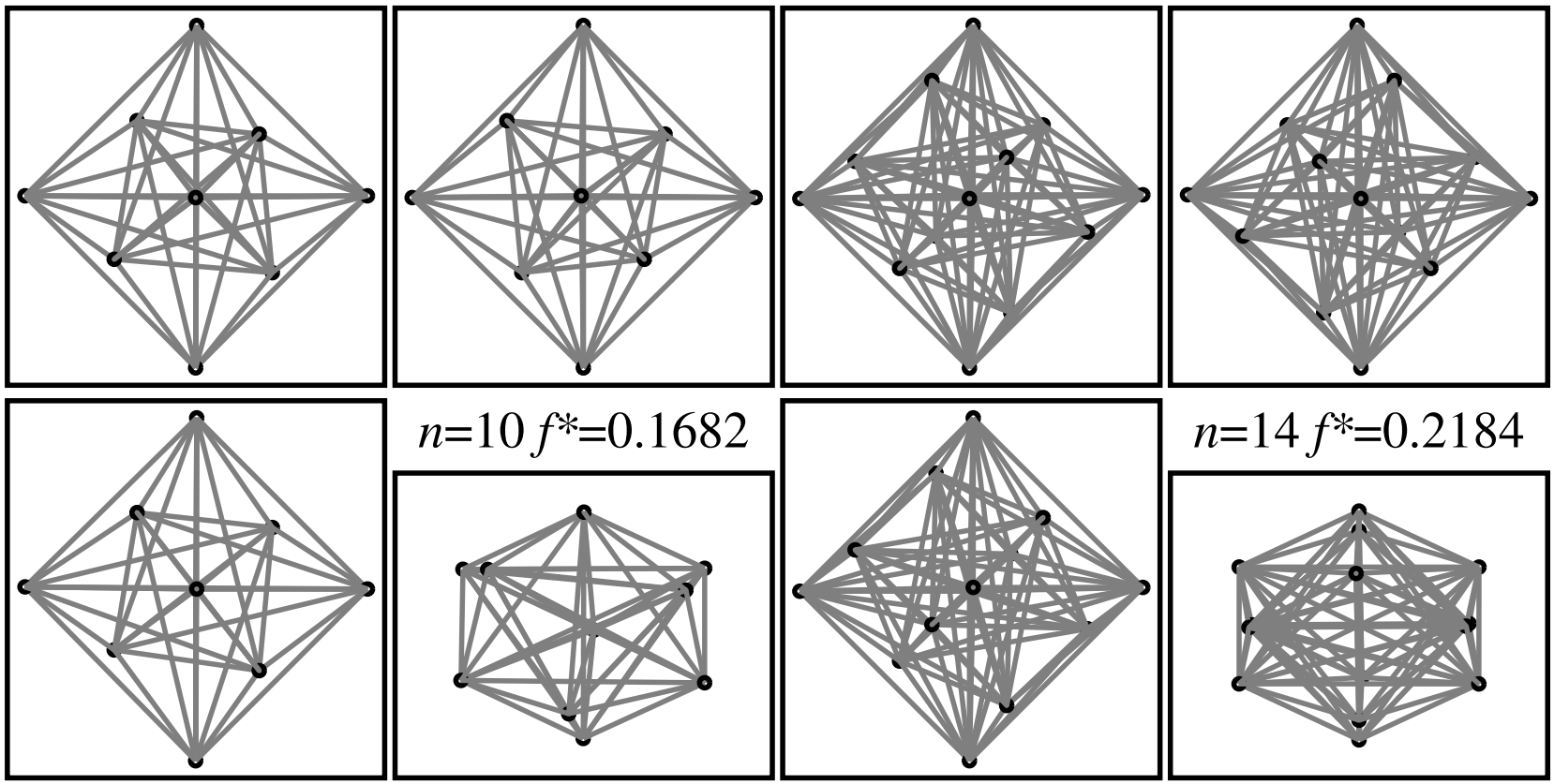
‘morseodes’



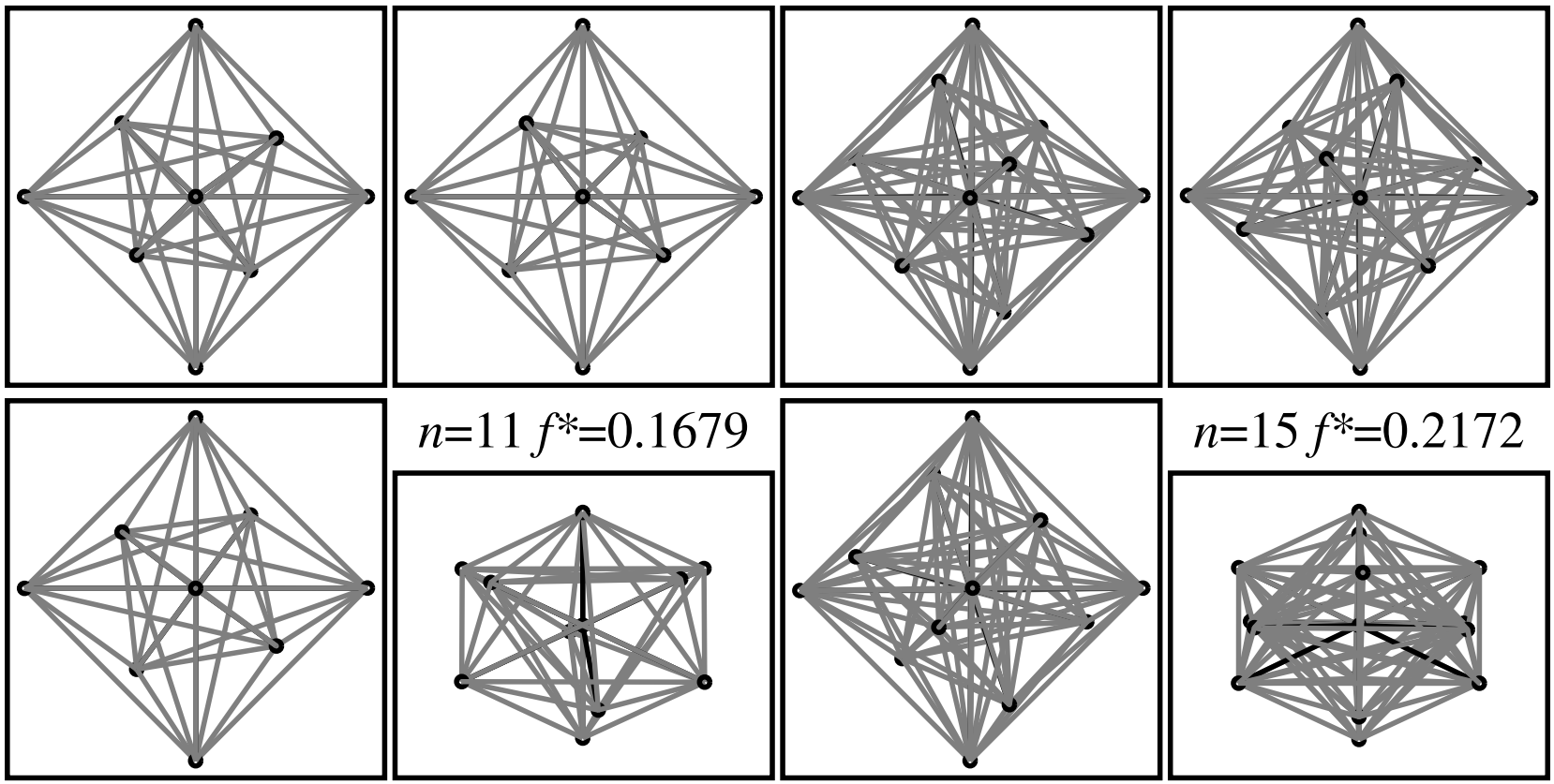
Images of pharmacological data, $m = 2$



Images of standard simplices, $m = 3$



Images of unit simplices, $m = 3$



Images of cubes, $m = 3$

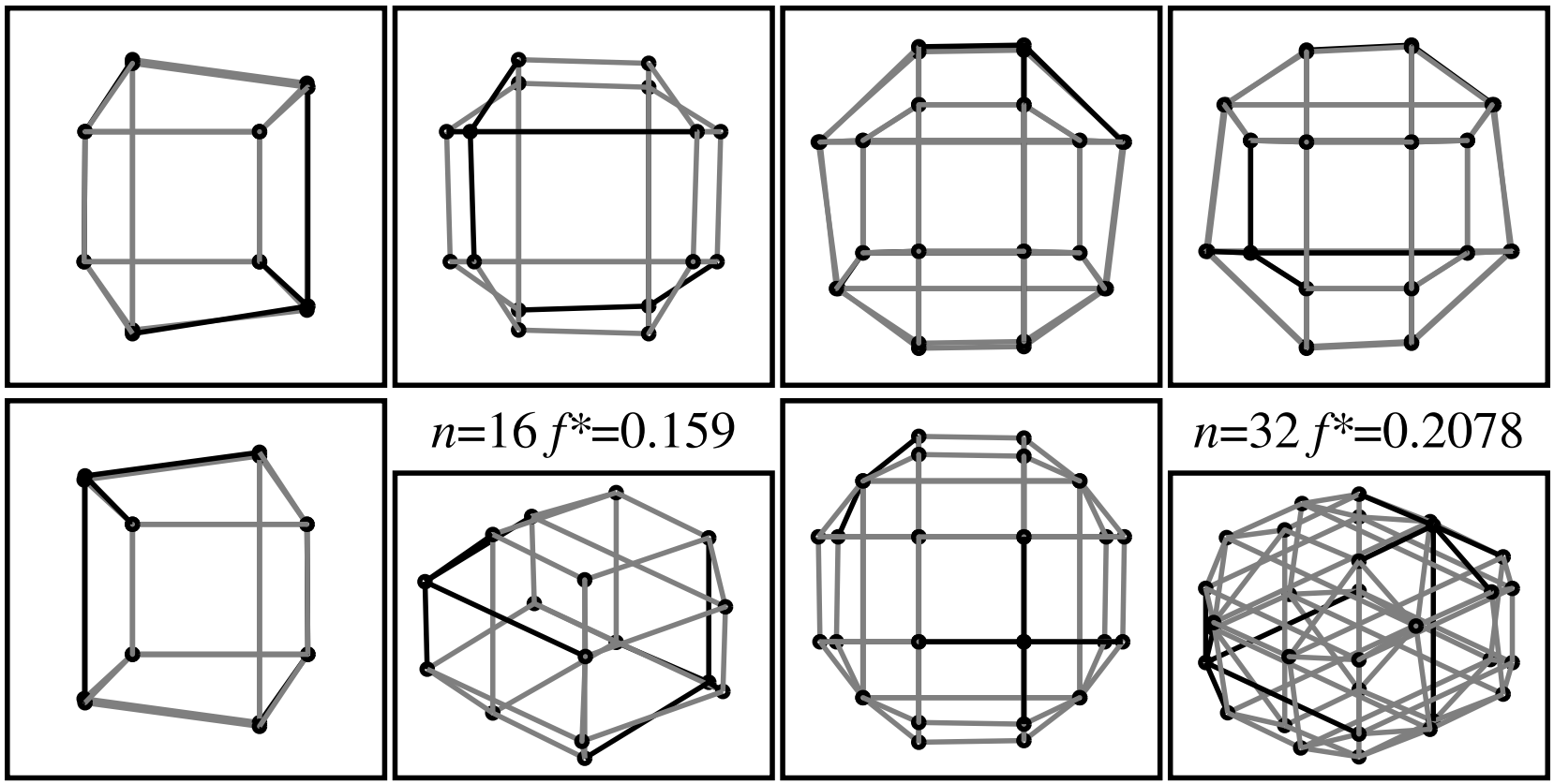


Image of the properties of human and zebrafish α_2-adrenoceptors, $m = 3$

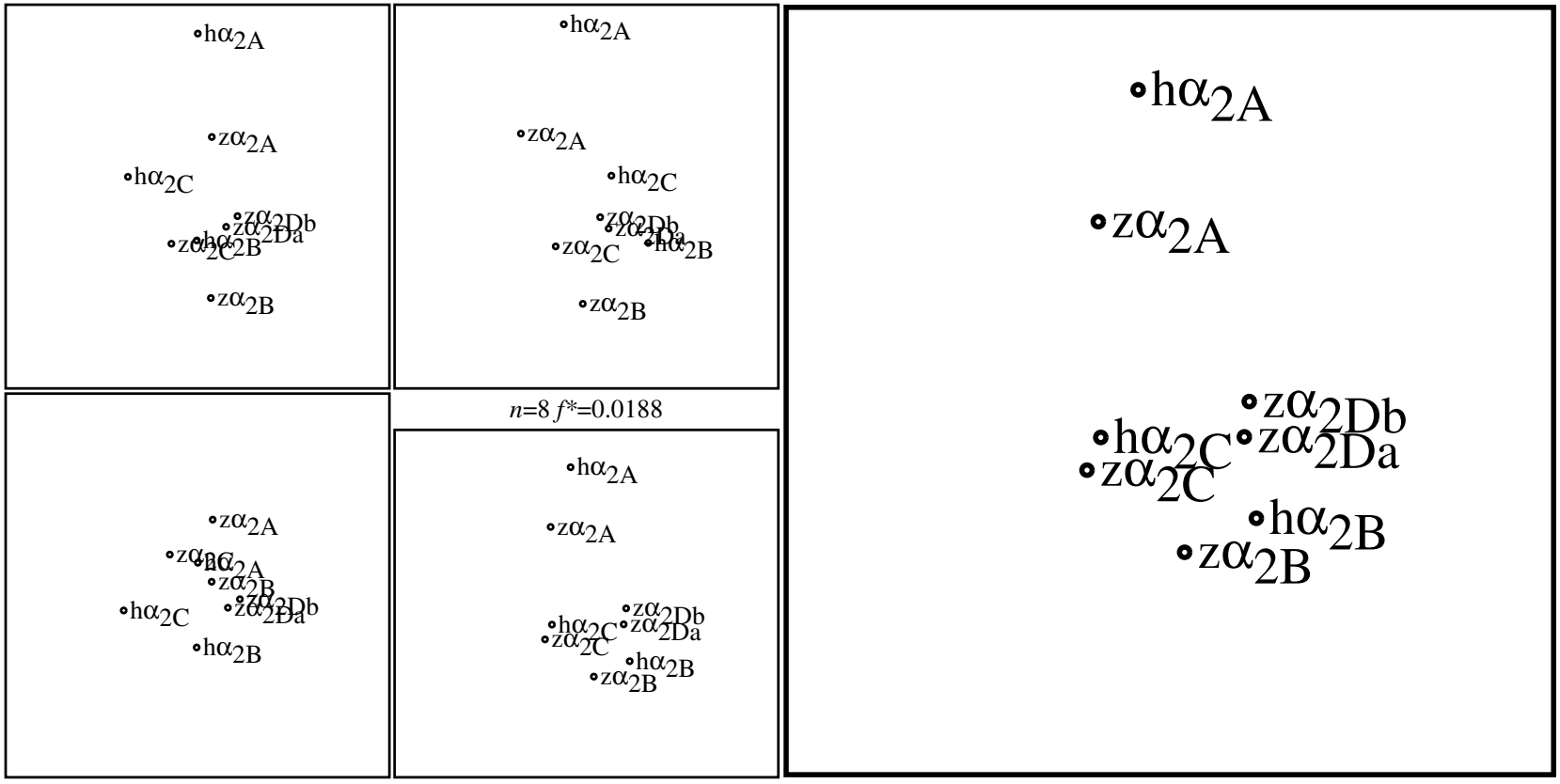


Image of the properties of 20 ligands binding human and zebrafish α_2-adrenoceptors, $m = 3$

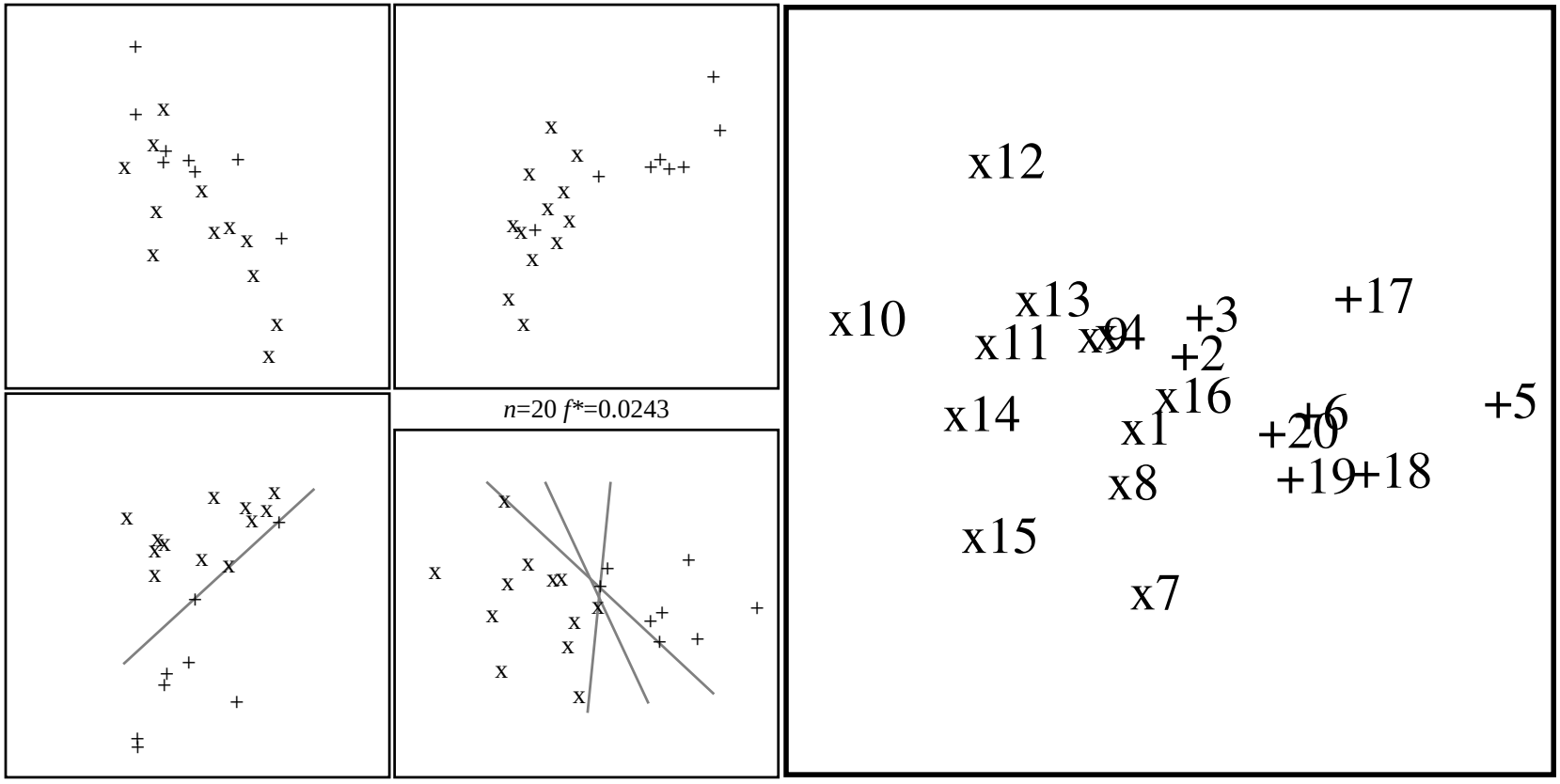


Image of the properties of human, rat, guinea pig and pig α_2-adrenoceptors, $m = 3$

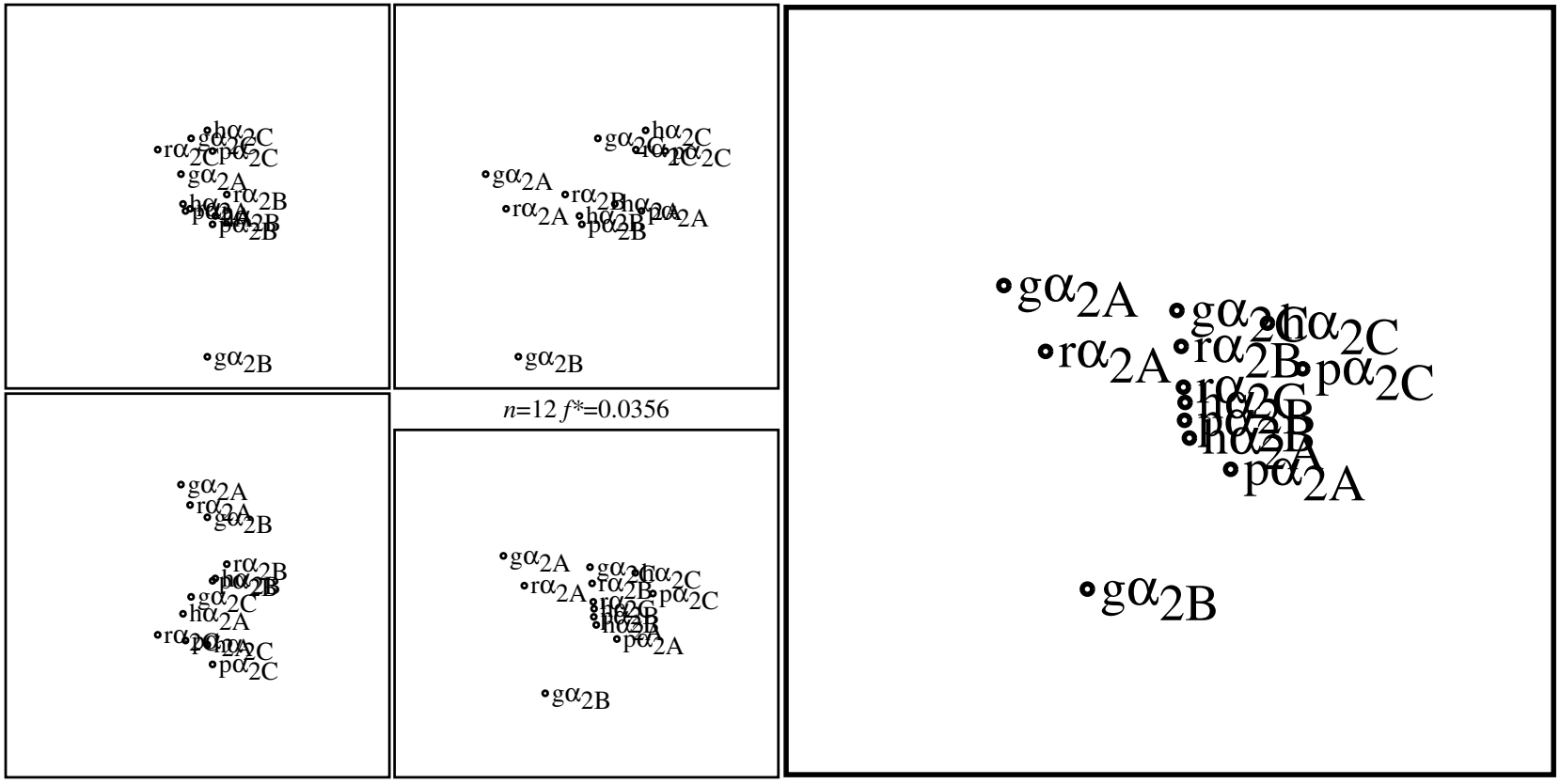


Image of the properties of wild type and mutant α_1-adrenoceptors, $m = 3$