For clustering vertices, the statistical decision theoretical problem statement is as follows:

1. A graph of a fixed size: $G_n = (V, E, Y)$:

   (a) V is the number of vertices.

   (b) E is the number of edges.

   (c) Y is vector labels $\{0, 1\}^n$. These are two distinct groups and we don't know them.

2. $SBM_n^k(\rho, \beta, ...)$ for k = 2:

   (a) $\rho \in \delta_2$ (2 simplex)

   (b) $\beta \in (0, 1)^{2x2}$

3. A = $\{y \in \{0, 1\}^n\}$. This is the set of cluster assignments.

4. $l : G_n \ x \ A \rightarrow R_+$; $l = \sum_{i=1}^{n} \Theta \{\hat{y}_i = y\}$

5. R = A x A x P (functional on true data, action class, model distribution)

1