

# A Fast, Consistent Kernel Two-Sample Test

A. Gretton, K. Fukumizu, Z. Harchaoui, B. Sriperumbudur

---

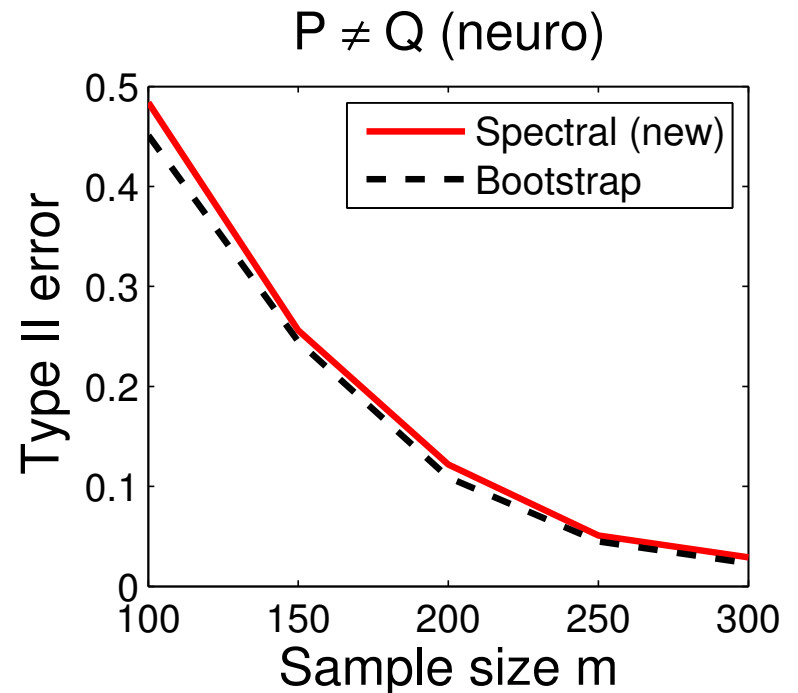
- Maximum mean discrepancy (MMD): distance between  $P$  and  $Q$

$$\text{MMD}(P, Q) := \|\mu_P - \mu_Q\|_{\mathcal{F}}^2$$

- Is  $\widehat{\text{MMD}}$  significantly  $> 0$ ?
- $P = Q$ , null distrib. of  $\widehat{\text{MMD}}$ :

$$m\widehat{\text{MMD}} \xrightarrow{D} \sum_{l=1}^{\infty} \lambda_l (z_l^2 - 2),$$

- $\lambda_l$  is  $l$ th eigenvalue of kernel  $\tilde{k}(x_i, x_j)$



Use Gram matrix spectrum for  $\hat{\lambda}_l$ : consistent test without bootstrap