Optimal Accuracy

[Bartlett & Traskin]

- usually, impossible to get perfect accuracy due to intrinsic noise or uncertainty
- Bayes optimal error = best possible error of any classifier
  - usually > 0
- can prove AdaBoost's classifier converges to Bayes optimal if:
  - enough data
  - run for many (but not too many) rounds
  - weak classifiers "sufficiently rich"
- "universally consistent"
- related results: [Jiang], [Lugosi & Vayatis], [Zhang & Yu], ...
- means:
  - AdaBoost can (theoretically) learn "optimally" even in noisy settings
  - but: does not explain why works when run for very many rounds