

# Variational Inference: (Non-exhaustive) Pros/Cons

- Pros:
  1. Deterministic algorithm---easy to tell when you've converged
  2. Embarrassingly parallel over documents
- Cons:
  1. Speed: makes many calls to transcendental functions
  2. Quality: fully factorized variational distribution is a questionable approximator
  3. Memory usage: requires  $O(MNK)$  to store the per-token variational distributions