Claims about algorithm 1:

- Useful when the optimization of $Q(\cdot\,;\cdot)$ is simpler than computing $l(\cdot)$

- To marginalize bad data, and increase marginalization monotonically.