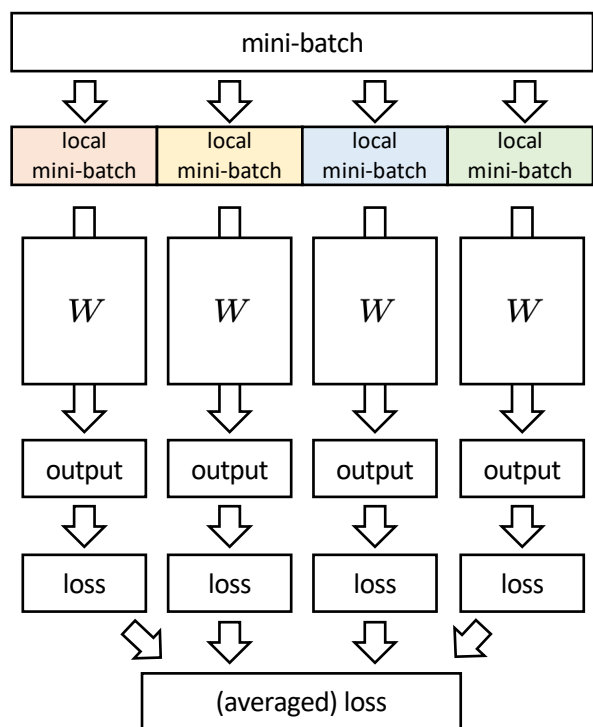


Data parallelism & MC sample parallelism for distributed VI training

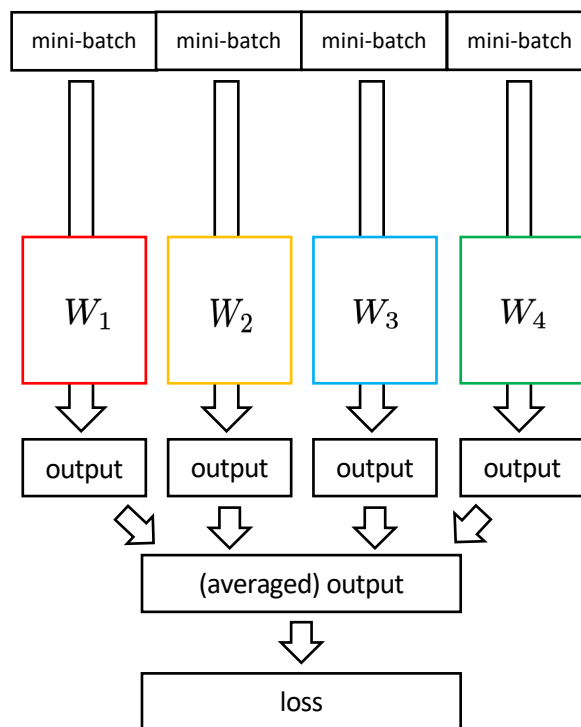
W : deterministic weights

Data parallelism
(among 4 processes)



$W_i \sim \text{posterior} (i = 1, 2, 3, 4)$

MC sample parallelism
(among 4 processes, 4 copies of mini-batch)



$W_i \sim \text{posterior} (i = 1, 2)$

Data & MC sample parallelism
(among 4 processes, 2 copies of mini-batch)

