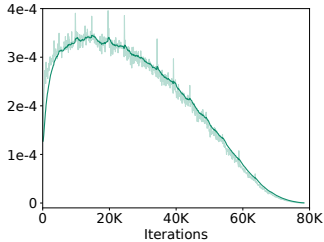
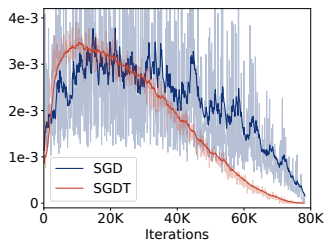


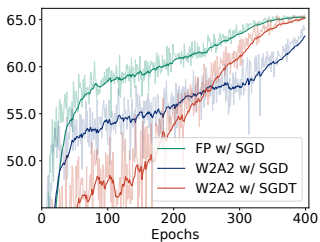
(a) LR decay for SGD.



(b) Avg. effective step sizes for FP weights (16<sup>th</sup> layer).



(c) Avg. effective step sizes for 2-bit weights (16<sup>th</sup> layer).



(d) Test accuracy (%).