Figure 1: Updated version of **Figure 1 (c)** with *AdamW* and *M-FAC without WD* (weight decay) graphs.
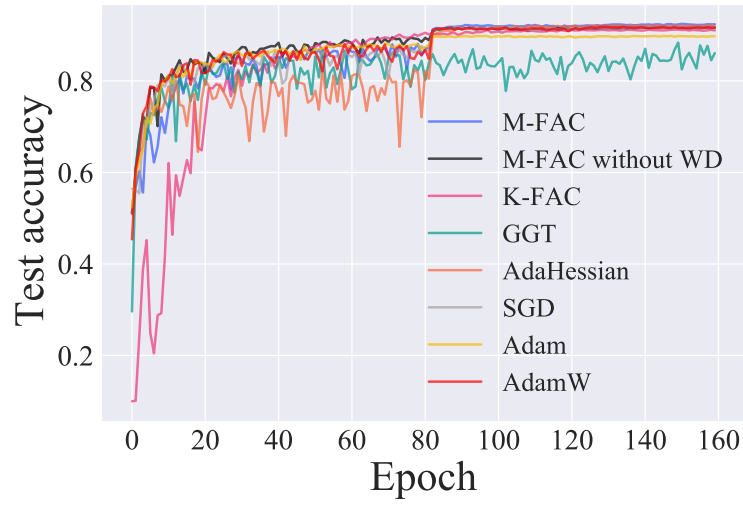

Figure 2: Updated version of **Figure 1 (d)** with *AdamW* and *M-FAC without WD* (weight decay) graphs.