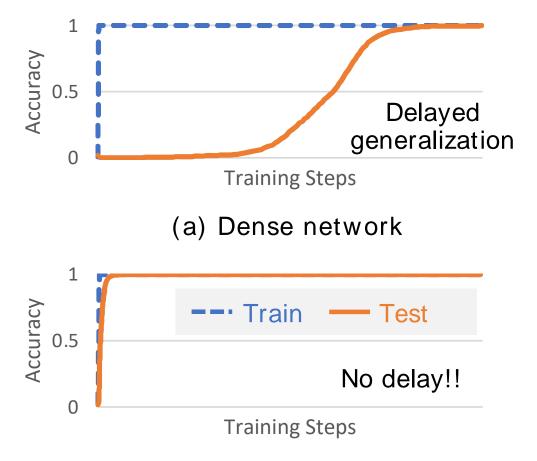
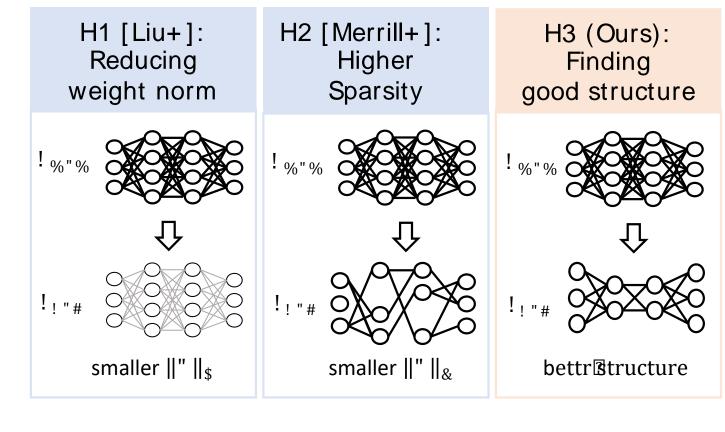
■ Observation (sec 3): Lottery ticket eliminates delayed generalization

What happen during delayed generalization? Why LT eliminate delayed generalization?





(b) Lottery Ticket (Good Subnetworks)