Test accuracy against learning rate η and weight decay d for different values of batch size s
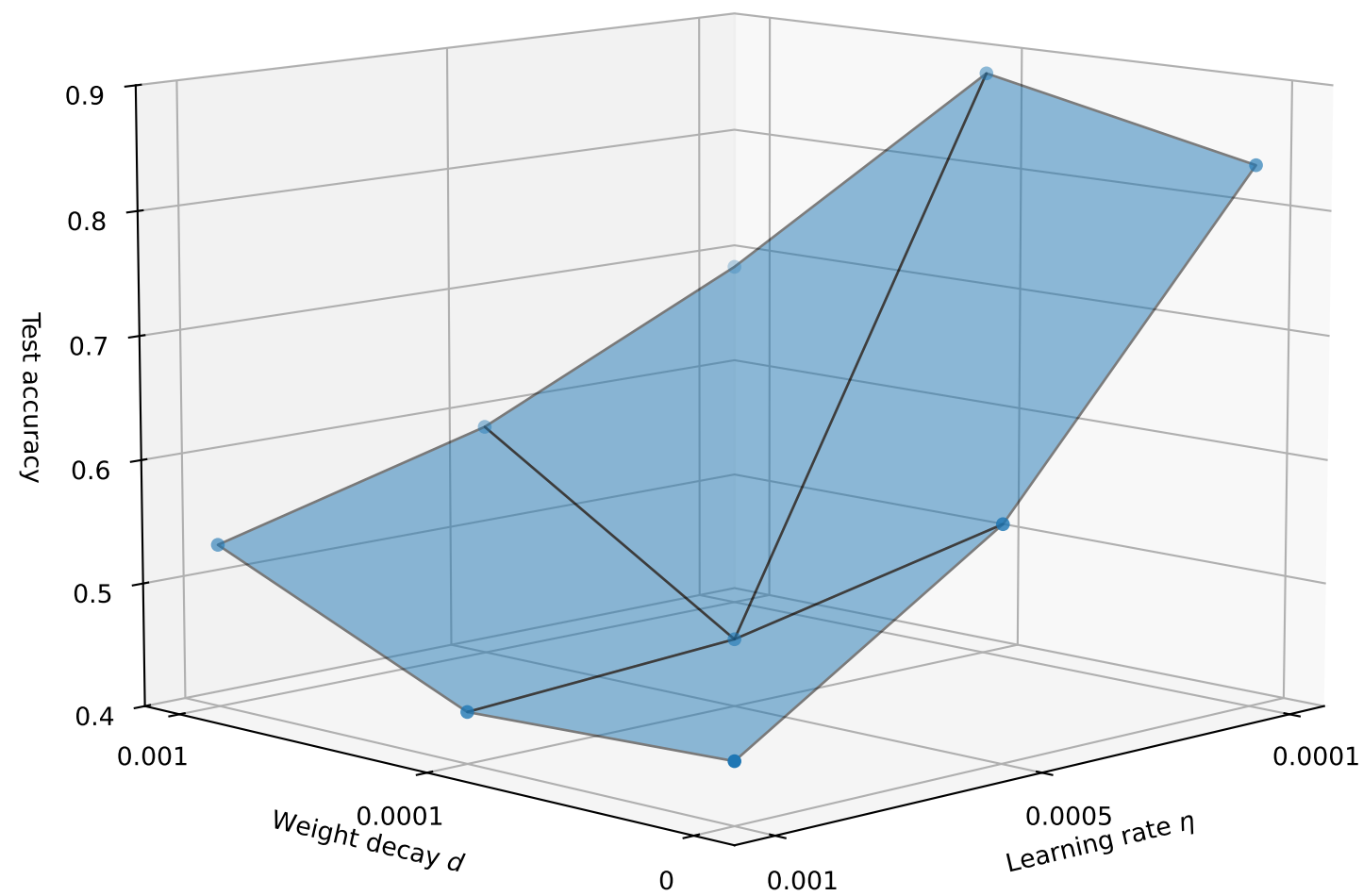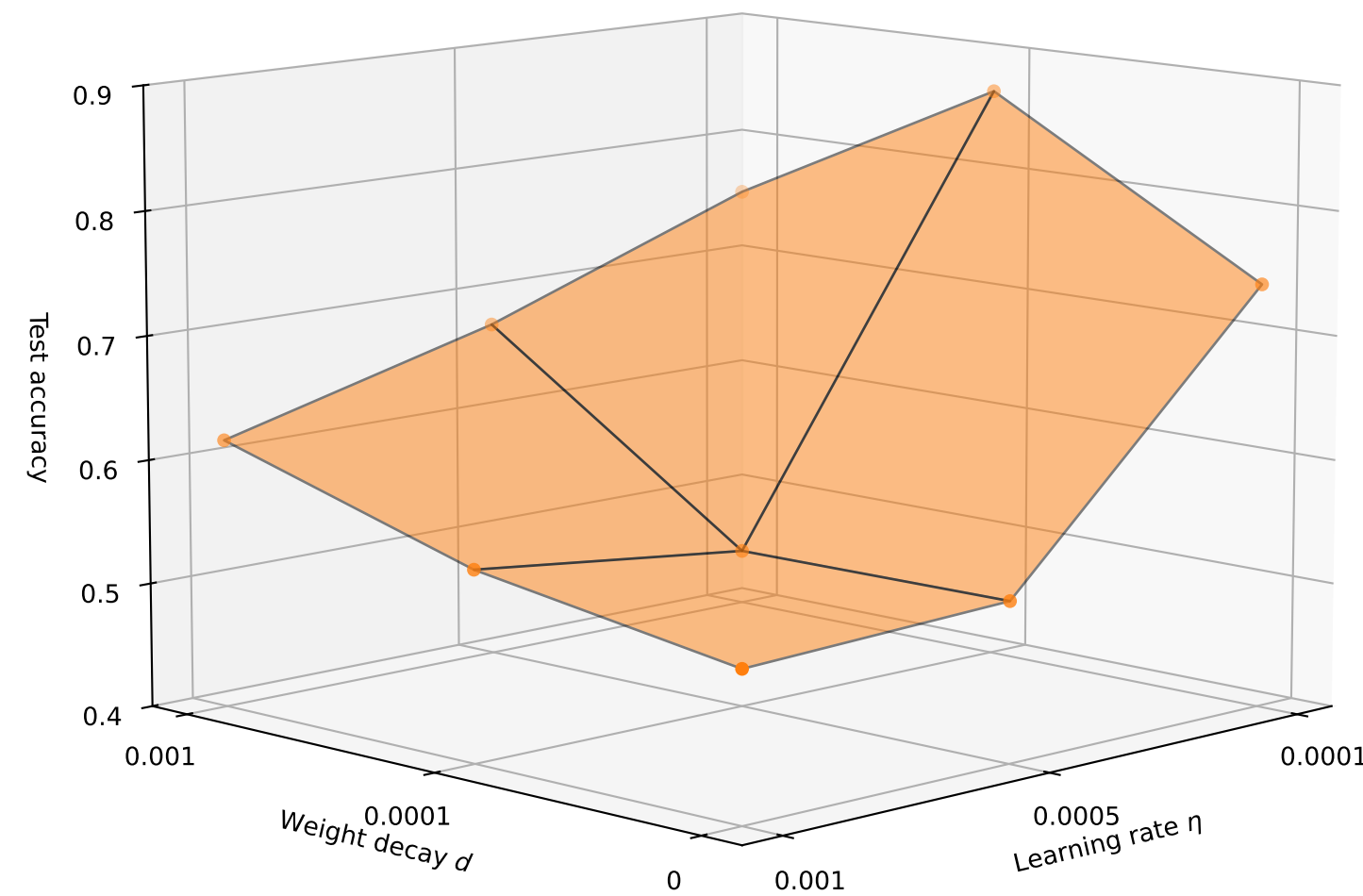
Batch size $s = 32$

Batch size $s = 64$

Batch size $s = 128$