

[HW4_prob1] Resnet-20 Post-training Quantization (10pts)

- Apply quantization for all the conv's weights by applying the functions in HW3_prob1's Resnet-20 by using following for loop.

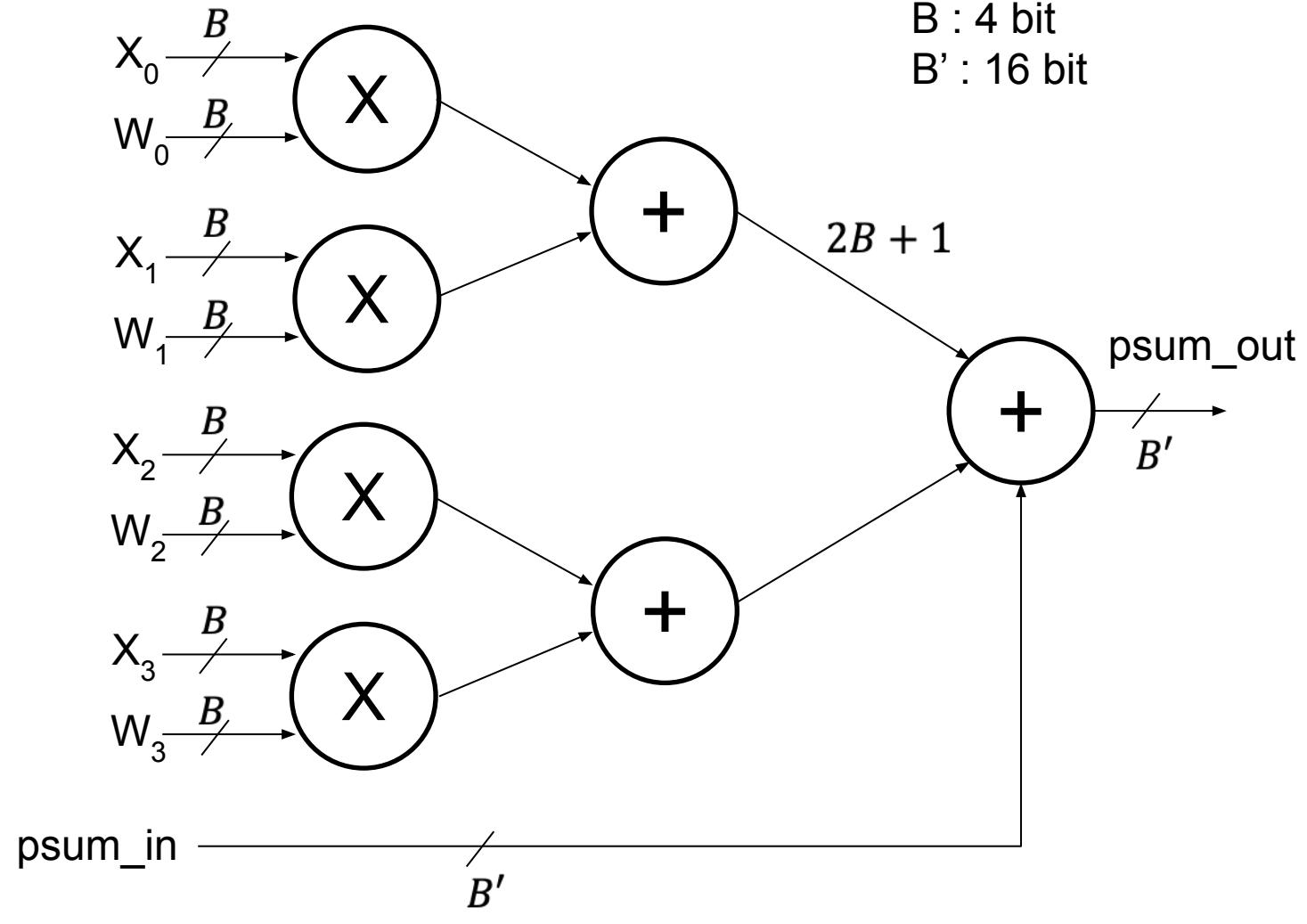
```
for layer in model.modules():
    if isinstance(layer, torch.nn.Conv2d):
```

- Try 4 bit and 8 bit quantizations
- Report your observation
- Download the notebook as pdf and submit it

[HW4_prob2] MAC Design (A)(10pts)

- Weight: 4-bit signed
- Activation: 4-bit **unsigned**
- psum / output: 16-bit
- Please fill out the empty part in hw/w3/prob3/verilog to support above number representation
- Need to create another “dec2bin” function in testbench for unsigned number (w_bin and x_bin)
- Need to modify “mac_predicted” function accordingly for sign * unsign multiplication
- Now, a_data.txt includes only positive numbers.
- Include your *.v files and capture the final waveform to show output

[HW4_prob3] MAC Design (B)(10pts)



- Now extend your previous solution to support the left schematic
- Again, your $psum_out$ bit precision is 16 bit
- Use the first four numbers in $a_data.txt$ as x_0, x_1, x_2, x_3 . Then, the next four numbers are x_0, x_1, x_2, x_3 for the following cycle.
- You need to modify both $_tb.v$ and other $*.v$ files accordingly.
- Include your $*.v$ files and capture the final waveform to show output
- all the multiplication and accumulation should happen in 1 cycle.