Deep-learning based approach for an OFDM System

Philip Lee Hann Yung

OFDM: Orthogonal Frequency Division Multiplexing

- A form of multicarrier modulation that is fundamental to LTE and Wi-Fi. Resilient to Inter-Symbol Interference (ISI).

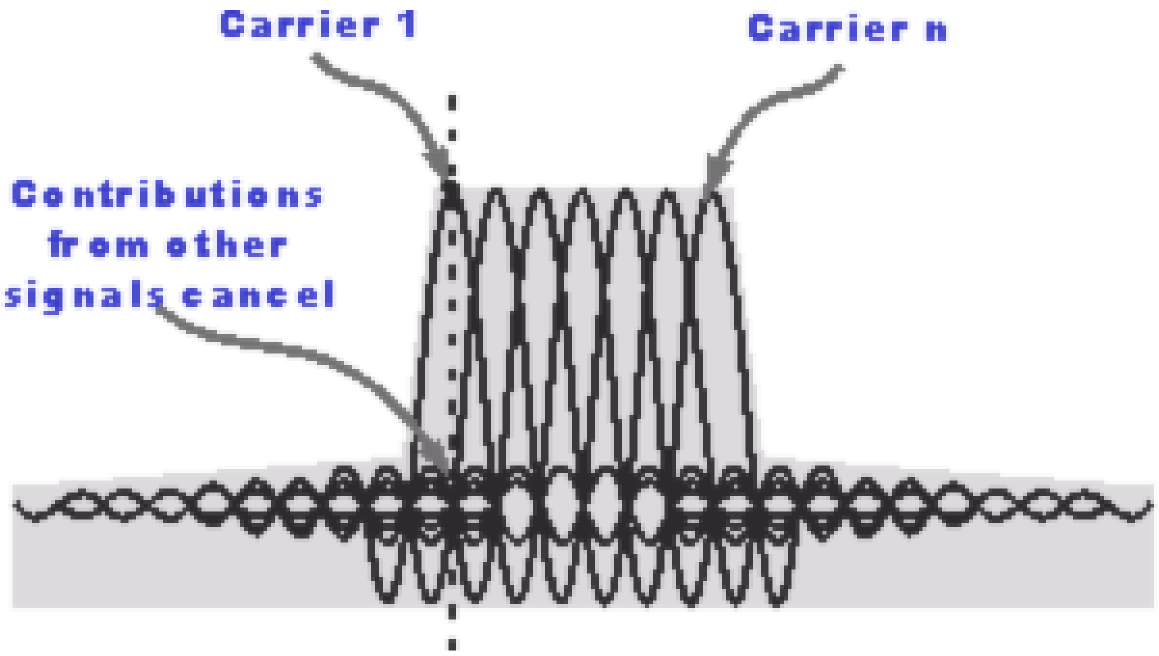


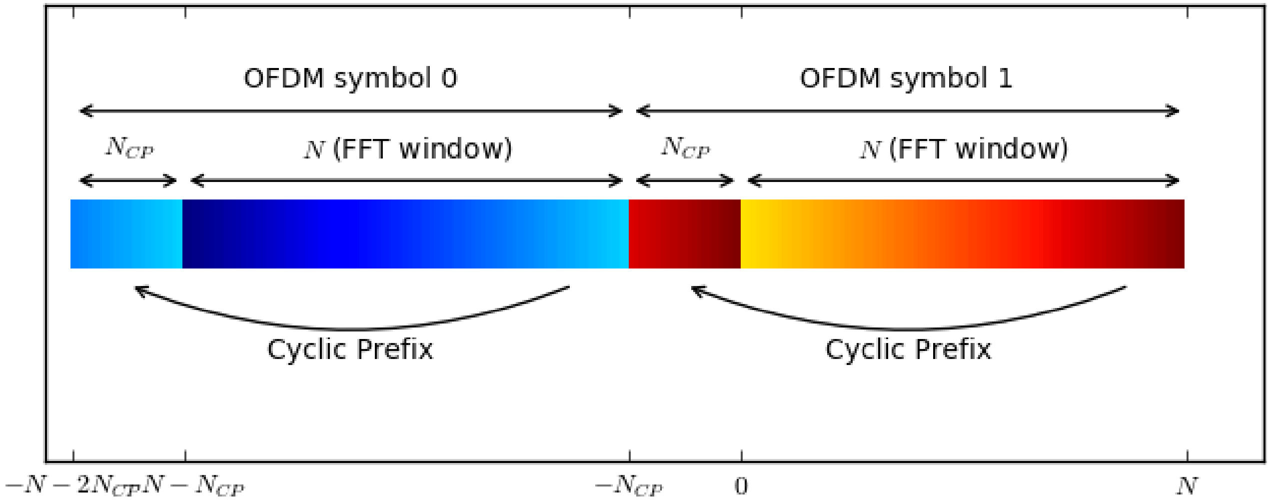
Image source: https://www.cablefree.net/wirelesstechnology/ofdm-introduction/

Motivation

Will a deep learning approach be able to recover data without CP effectively?

Role of Cyclic Prefix (CP):

- Turns linear convolution into circular convolution.
- Inserted as a guard band to protect against Inter-Symbol Interference (ISI).

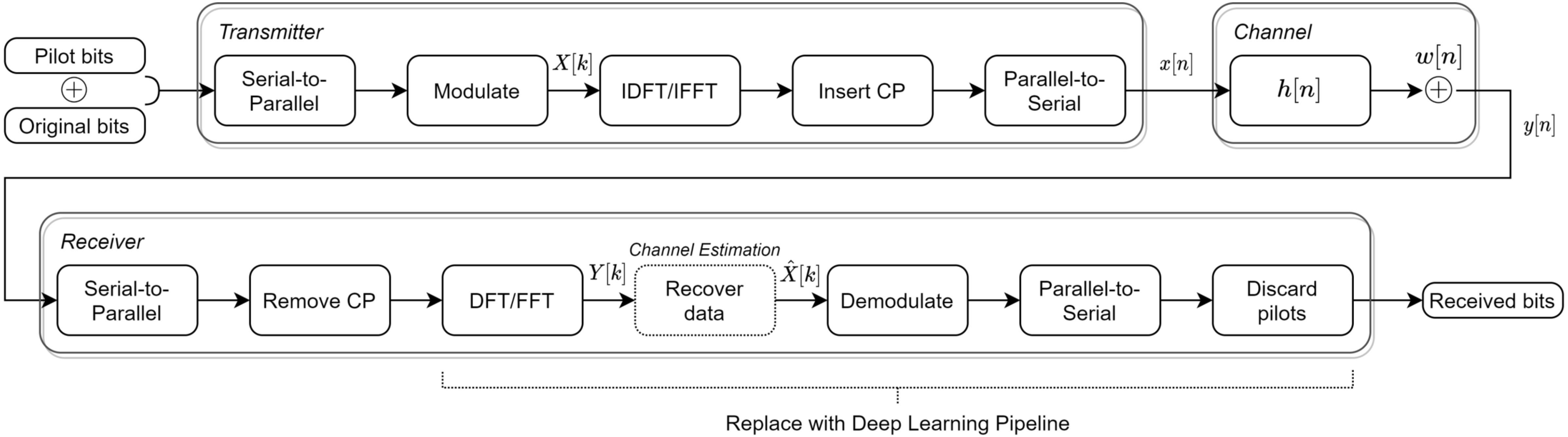


Disadvantages:

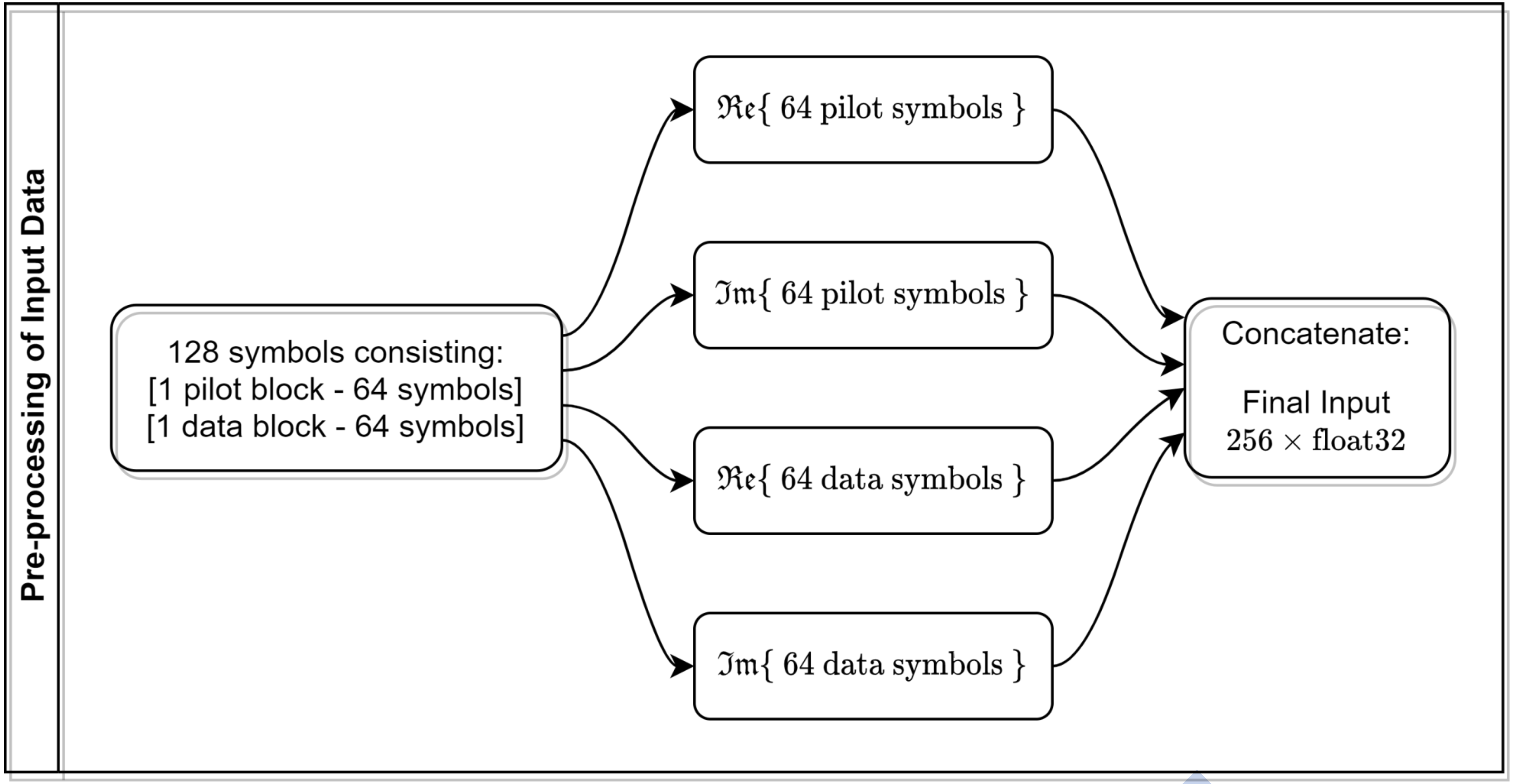
- Additional power consumption.
- Reduced overall data rate due to CP overhead.
  - ➢ Removing CP will result in more challenging data recovery.

Image source: https://dspillustrations.com/pages/posts/misc/the-cyclic-prefix-cp-in-ofdm.html

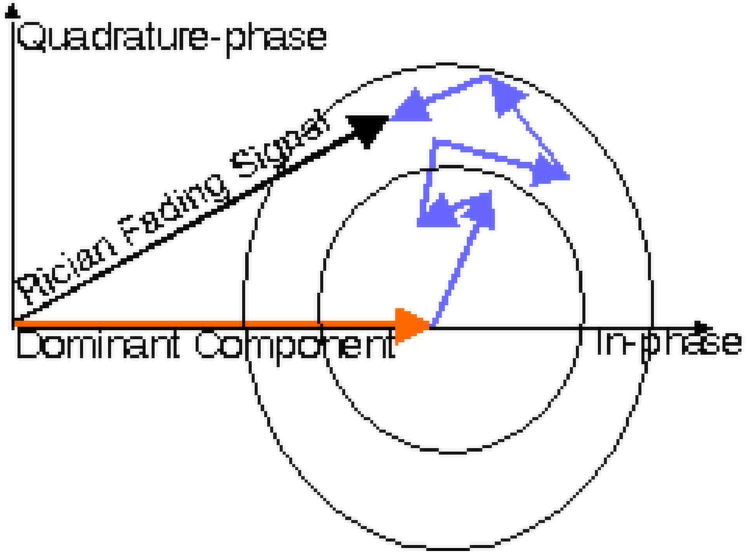
Typical OFDM System Structure



Inputs Pre-processing



Methodology



- Rician fading channel
- Monte Carlo simulations are performed on OFDM with SNR 20 dB ← no effect
- Training Set Diversity
  i. Rician K-factors : -40 to 20 dB (higher means larger LOS path)
  ii. Channel taps : Length of 3-10 (number of reflection paths)
- Baseline comparisons: LS and LMMSE (two conventional channel estimation methods)

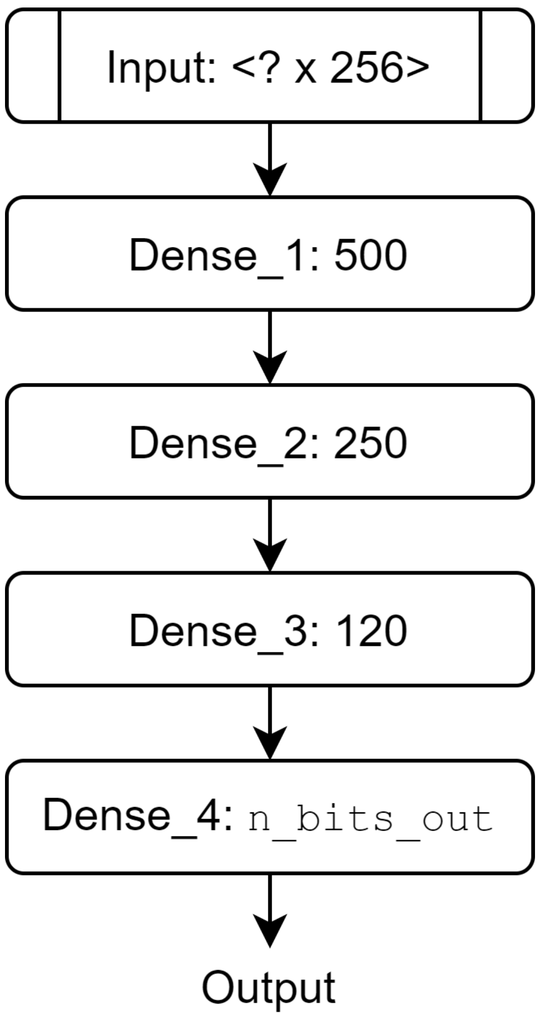
Image source: http://www.wirelesscommunication.nl/reference/chaptr03/ricepdf/rice.htm

Initial Implementation

- Loss function : Binary cross-entropy
- Evaluation metric : BER (lower is better)
- Trainable parameters : 300k
- Epochs trained : 150
- Only predict 64 of 128 bits

References:

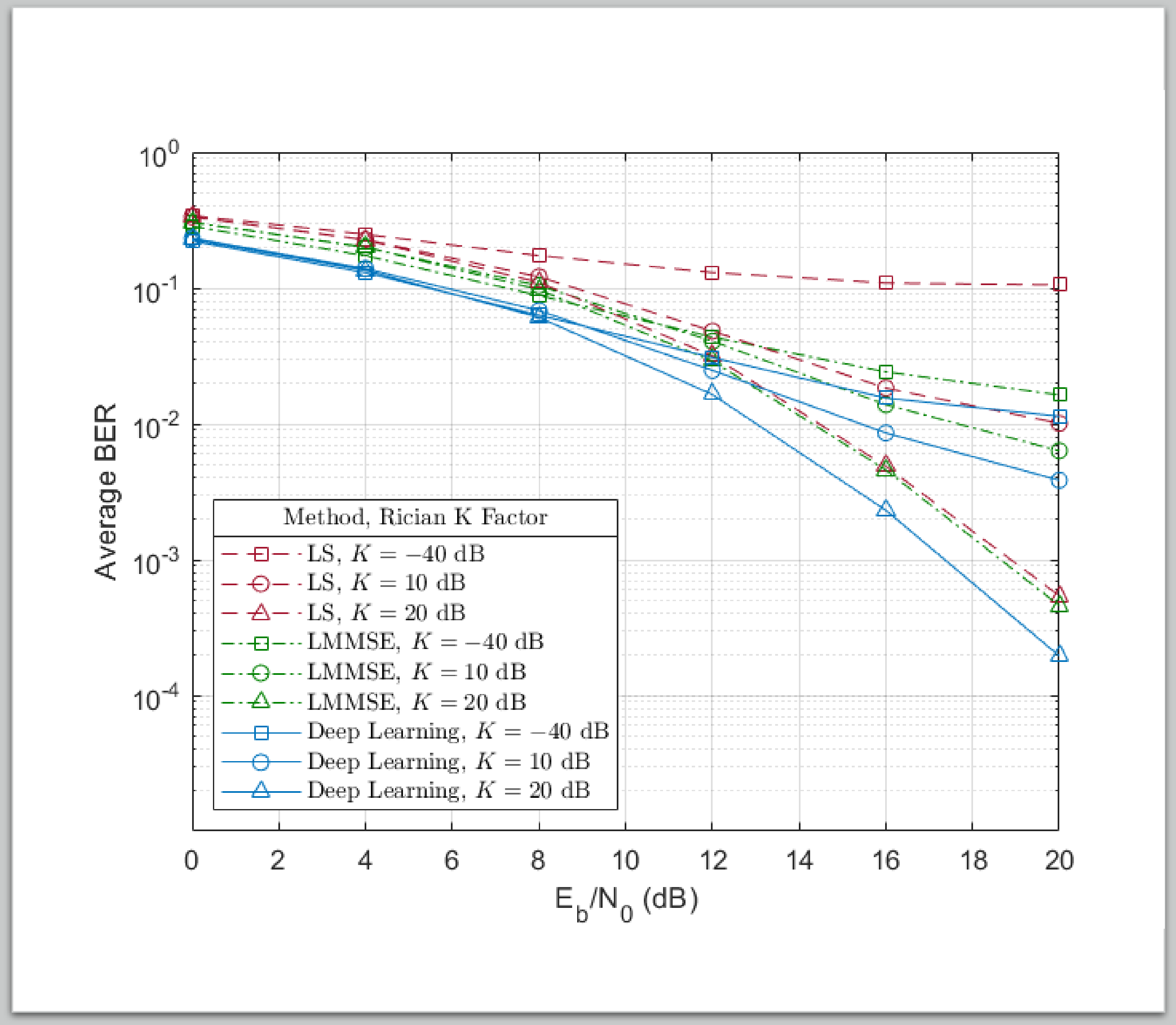
1) https://github.com/haoyye/OFDM_DNN
2) H. Ye, G. Y. Li, and B.-H. Juang, “Power of deep learning for channel estimation and signal detection in OFDM systems,” IEEE wirel. commun. lett., vol. 7, no. 1, pp. 114–117, 2018.



Model 1:
OFDM without CP

- It performs the best but there was a compromise of only 50% data rate (64 of 128 bits prediction)

Note: DL method is the worst for OFDM with CP.



Conflict with Original Motivation

Why wasn't 128 total bits predicted instead?

- BER is very high (almost unusable)

How 64/128 bits was chosen? Currently, BER is not bad …

- When the model was used to predict a smaller number of bits e.g. 8, 16 bits, the BER was lower but there is high data wastage.

Inspiration: Siamese Networks

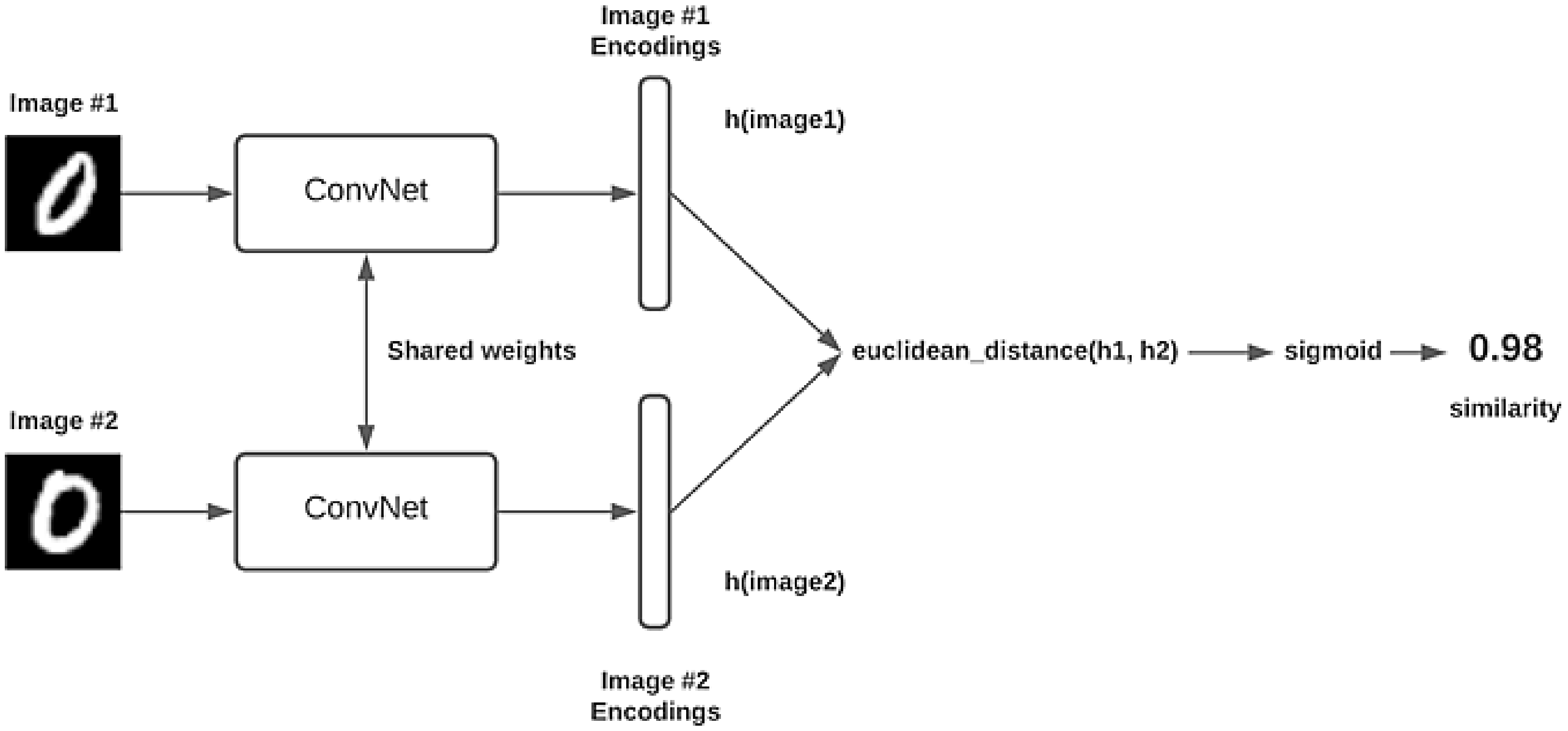


Image source: https://www.pyimagesearch.com/2020/11/30/siamese-networks-with-keras-tensorflow-and-deep-learning/

Improved Model

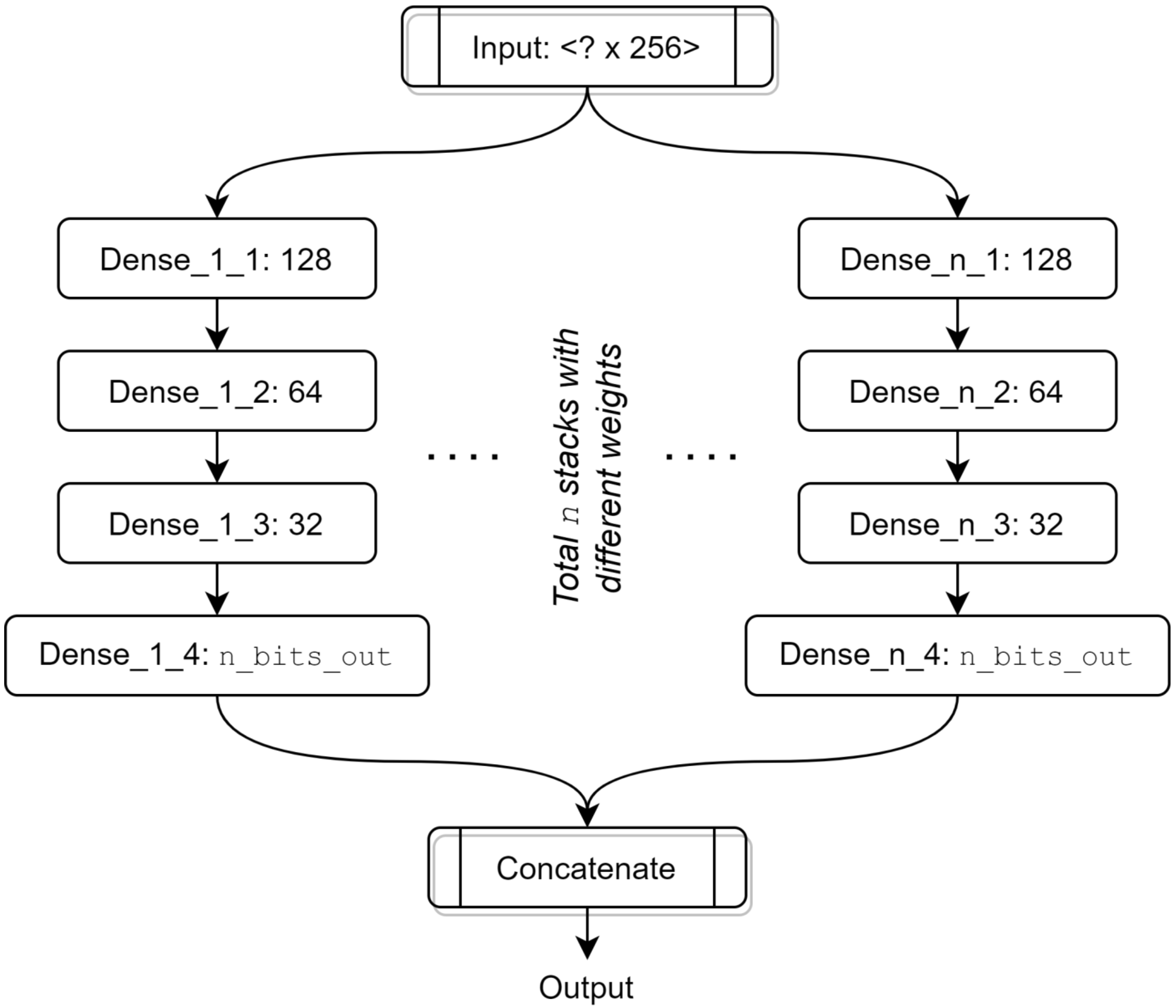
`n_bits_out` = 8 & 16 stacks:

- 700k trainable parameters
- 67% savings in training time

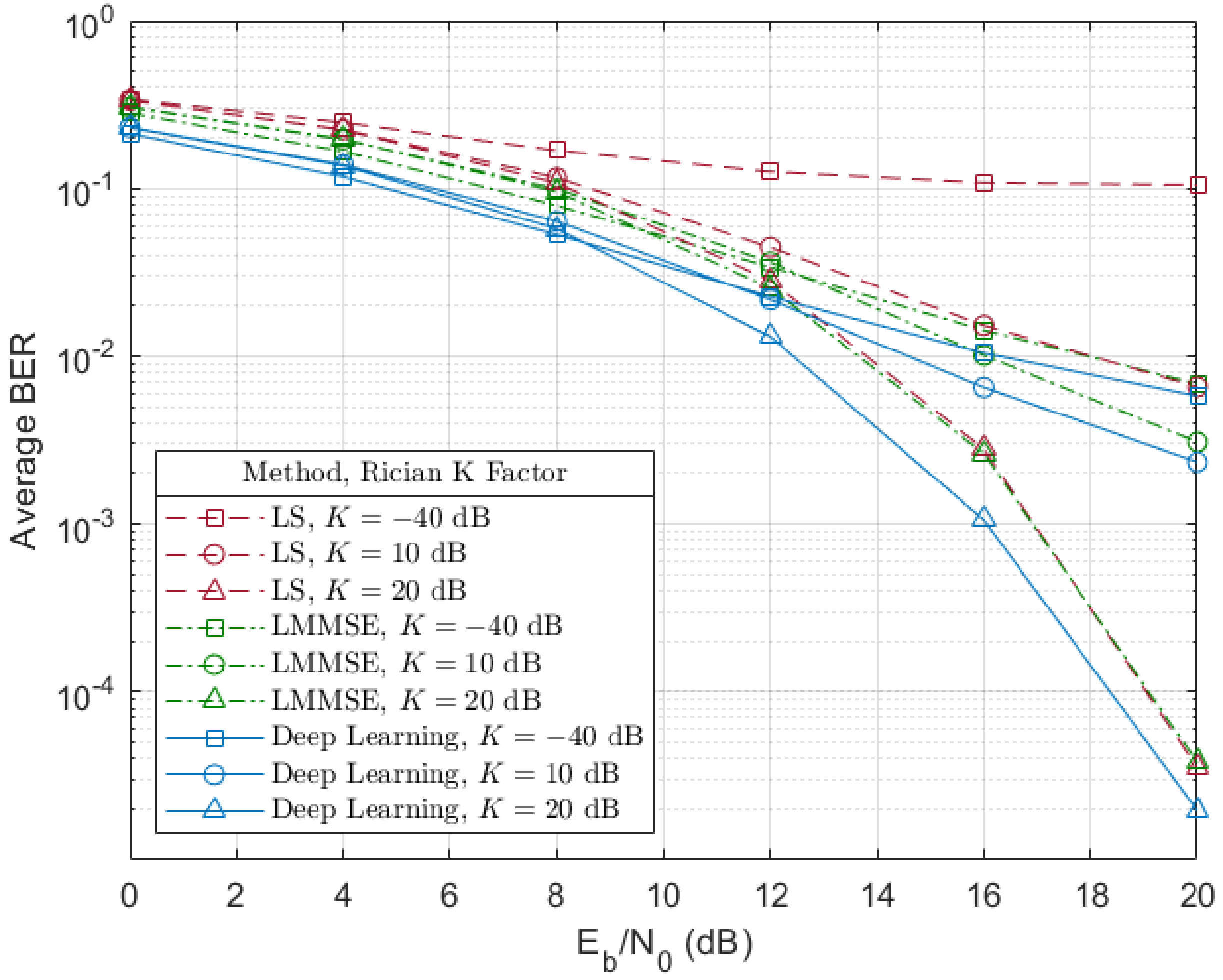
Most importantly:

- All 128 bits can be predicted!
- Improved BER

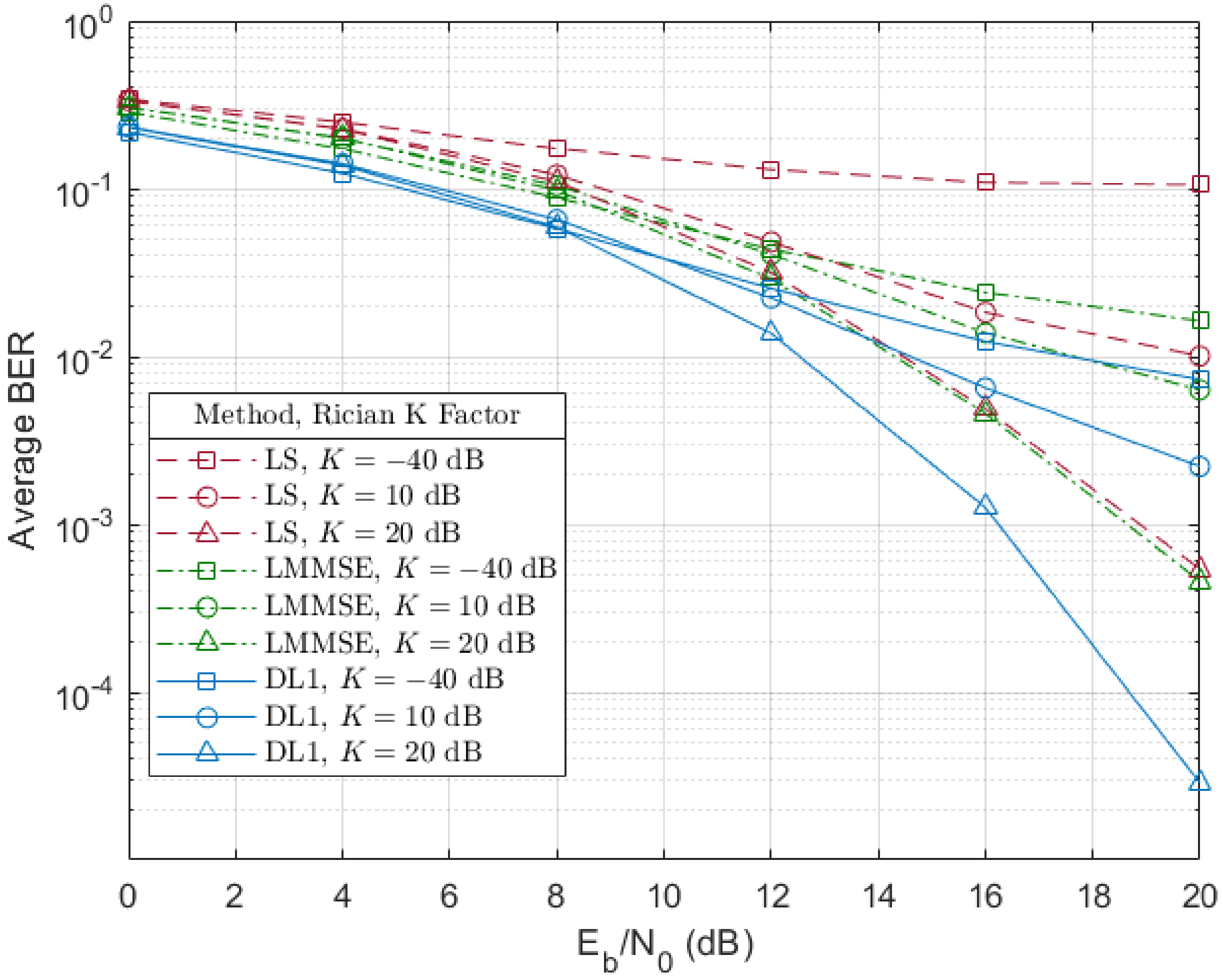
Note: Model demonstration is on 10 channel taps



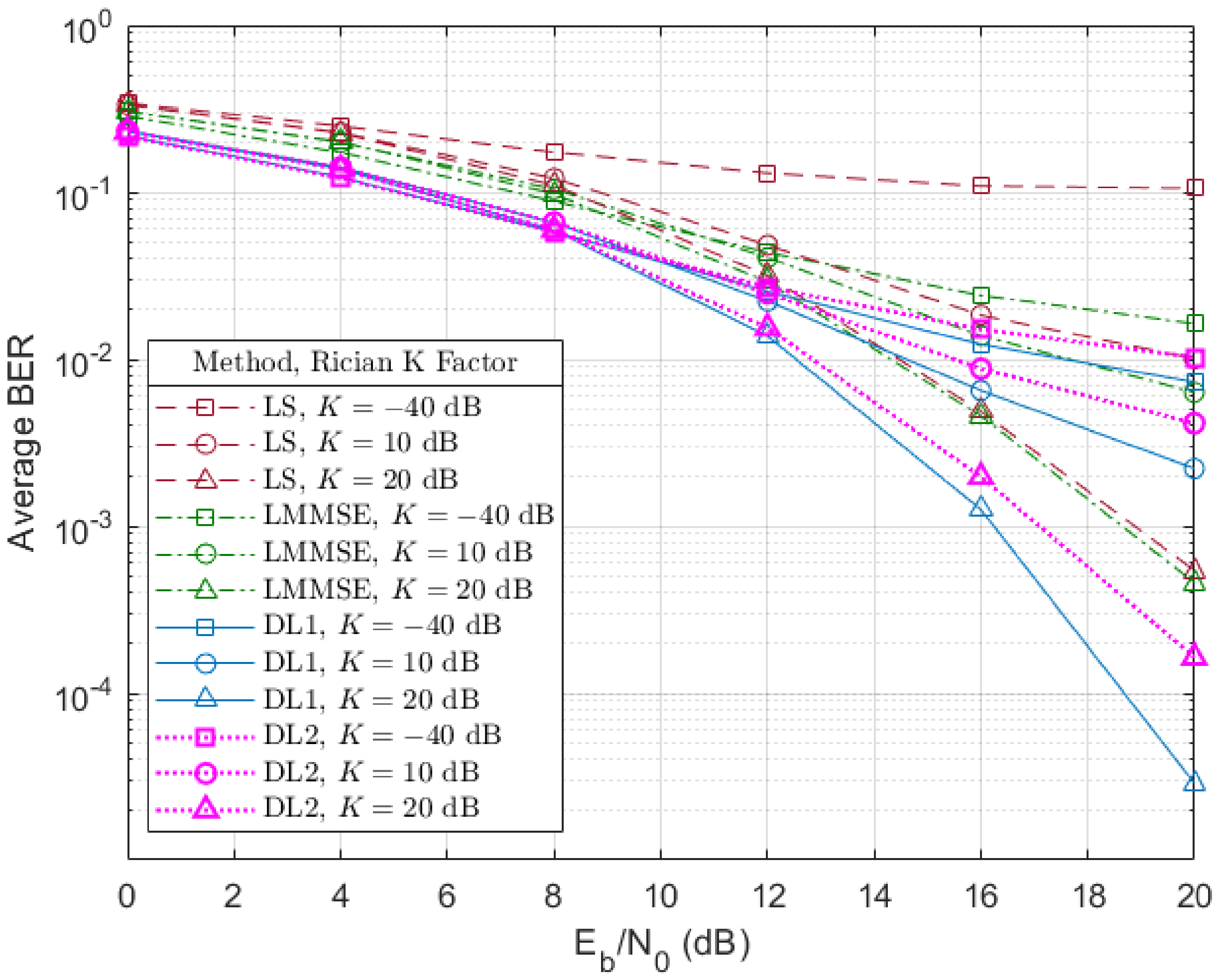
Model 2A: OFDM with CP



Model 2B: OFDM without CP



DL1: Model 2B
DL2: Model 2A



Conclusion

- Lower BER
  - DL method is able to recover data better than other methods.
- One-method-for-all
  - Using the same trained model, DL1 has lower BERs across all 3-10 channel taps, without knowing second-order channel statistics
- Robustness
  - DL2 shows that BER only suffers a little when using a model trained on a normal situation with CP.

Future Development

- System/commercial device can auto-train to perform optimal data recovery.
- The only drawback might be the time taken to perform data recovery, but this may be improved by:
    - Direct hardware implementation.
    - Performing data recovery in batches and in parallel.

Reference: S. Oh et al., “Energy-efficient Mott activation neuron for full-hardware implementation of neural networks,” Nat Nanotechnol., 2021.


Thank you!