MADD 2020

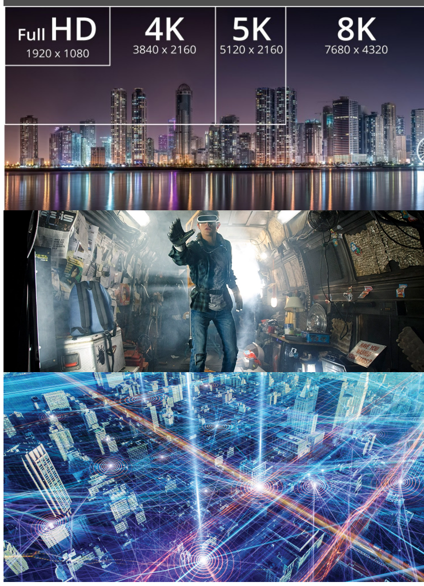
Deep Learning: Not a Simple Hammer for Massive MIMO Wireless Communication Systems

Zhi Ding
UC Davis
zding@ucdavis.edu

Outlines

- Introduction
- Massive MIMO: Promises and Challenges
- Deep Learning and CSI Estimation
- CSI Correlations in Massive MIMO
- Exploit CSI Characteristics in Deep Learning-
- Open Issues

Demands for Faster & Better Communication Networks

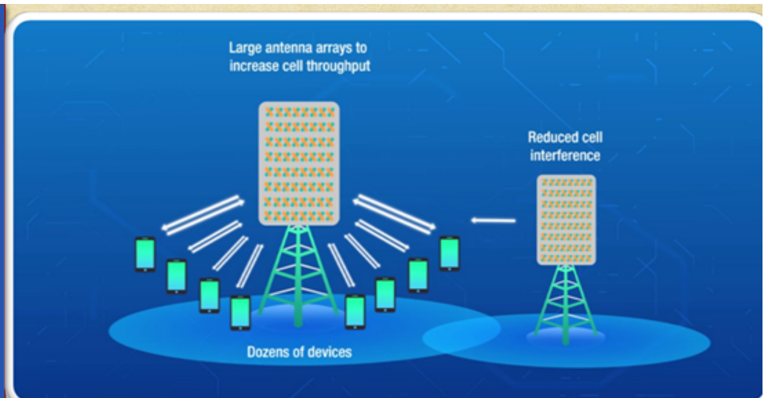


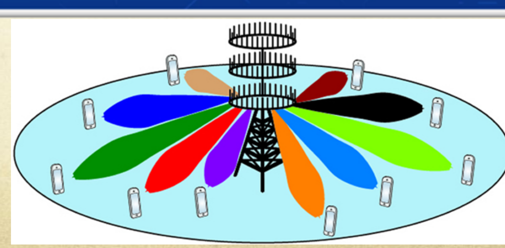
- Ultra high data rate
- Ultra low delay
- Massive number of connected devices

MASSIVE MIMO Promises:

- Benefits
  - Increased capacity
  - Improved robustness
  - Reduced latency
  - Higher efficiency

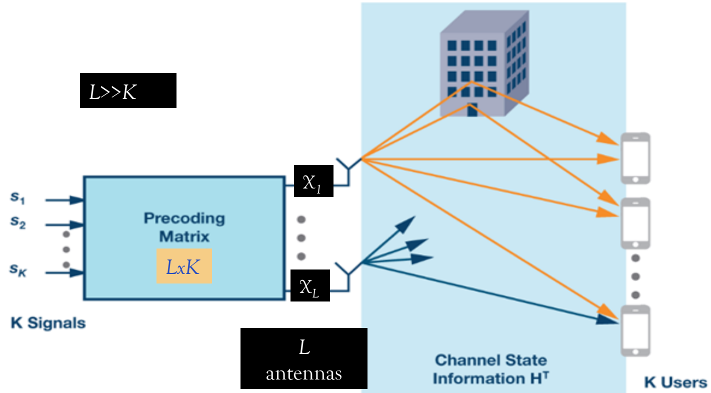
[1]Erik G. Larsson, et al. "Massive MIMO for Next Generation Wireless Systems." IEEE Communications Magazine 52.2 (2014): 186-195.





MASSIVE MIMO Downlink Challenges

- Accurate downlink CSI for precoding and resource management
- TDD: reciprocity
- FDD: UE feedback
- High CSI feedback load
  - Many antennas
  - Many users
  - Wide bandwidth



Precoding Type	
Maximum Ratio (MR)	$\mathbf{x} = H^{*}\mathbf{s}$
Zero Forcing (ZF)	$\mathbf{x} = H^{*}(H^{T}H^{*})^{-1}\mathbf{s}$
NMSE or RZF	$\mathbf{x} = H^{*}(H^{T}H^{*} + \beta I)^{-1}\mathbf{s}$

[1] https://www.analog.com/en/analog-dialogue/articles/massive-mimo-and-beamforming-the-signal-processing-behind-the-5g-buzzwords.html

Efficient CSI feedback

- Key FDD challenge: reducing massive MIMO CSI feedback
- Existing CSI feedbacks
  - Quantized CSI feedback
    - Large codebook size
    - Inefficient
  - Compressive sensing (CS)-based
    - Delay from iterative algorithms
    - Scalability

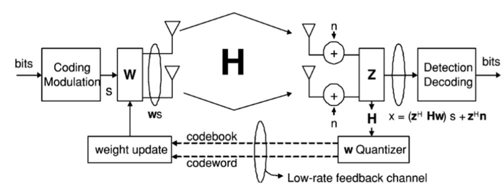
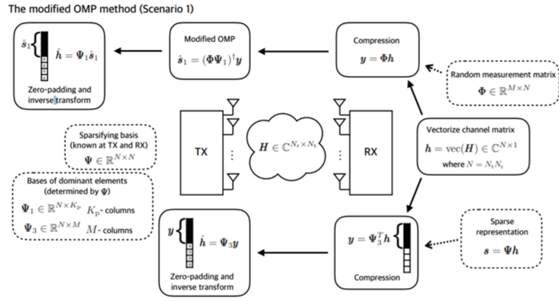


Fig. 1. MIMO beamforming and combining system with feedback and quantization.

The modified OMP method (Scenario 1)

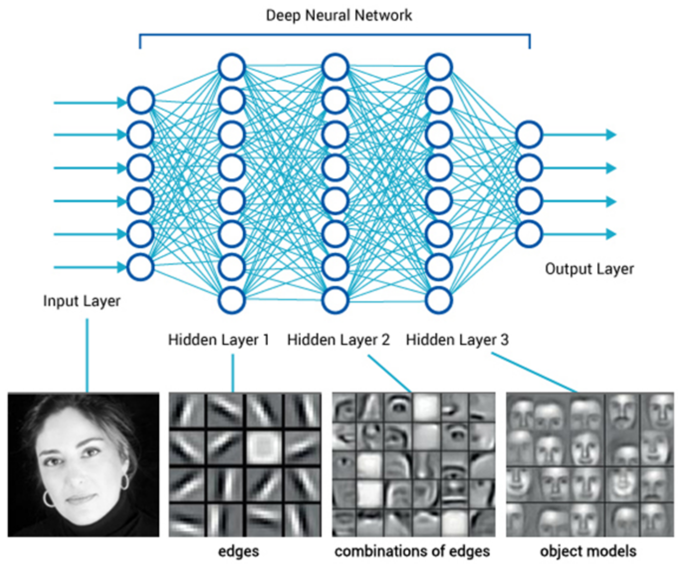


The dimensionality reduction method (Scenario 2)

[1] B. Mondal and R. W. Heath, "Channel Adaptive Quantization for Limited Feedback MIMO Beamforming Systems," in *IEEE Transactions on Signal Processing*, vol. 54, no. 1 pp. 4717-4729, Dec. 2006.
[2] Min Soo Sim, et al. "Compressed channel feedback for correlated massive MIMO systems." *Journal of Communications and Networks* 18.1 (2016): 95-104

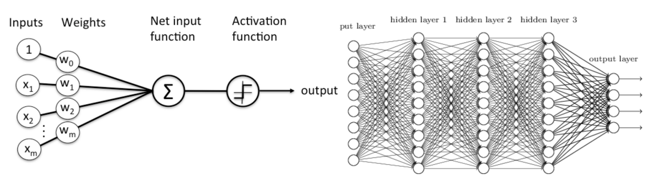
Deep Learning (DL)

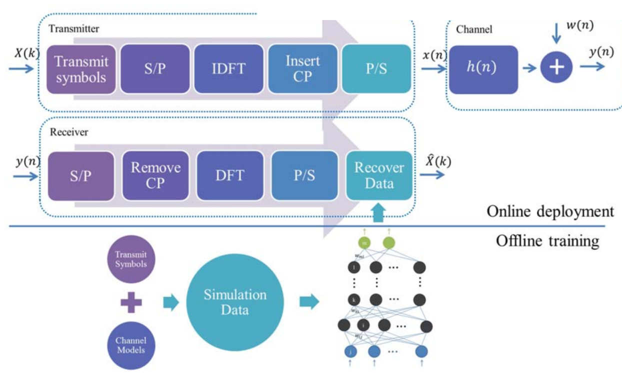
- Successes in CV, NLP, decision making ...
  - Alleviate need for feature engineering
  - Broadly adaptable to new problems
- DL in wireless: Coding, CSI estimation, detection, modulation classification, etc.



Deep Neural Network

- Perceptron
  - Neuron with adjustable weights and an activation function
- DNN
  - Multilayer perceptron
  - Back propagation
- Signal detection applications





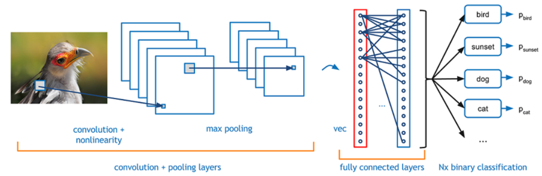
[1] Hao Ye, Geoffrey Ye Li, and Biing-Hwang Juang. "Power of deep learning for channel estimation and signal detection in OFDM systems." IEEE Wireless Communications Letters 7.1 (2018): 114-117.

Convolutional Neural Network (CNN)

- Convolutional layers + pooling layers + fully connected layers
- Take advantage of the 2D structure
- Fewer parameters

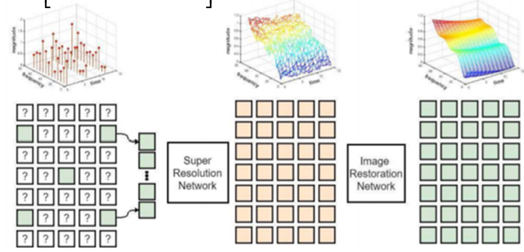
- Image recognition successes
- CSI matrix vs 2-D image (similarity)
- CSI estimation and feedback

(image compression?)



$$H=\begin{bmatrix} h_{1,1} & \cdots & h_{1,K} \\ \vdots & \cdots & \vdots \\ h_{L,1} & \cdots & h_{L,K} \end{bmatrix}$$

CSI-matrix in massive MIMO

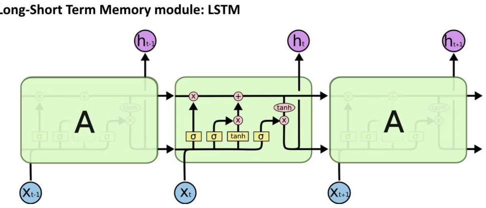


[1] M. Soltani, V. Pourahmadi, A. Mirzaei and H. Sheikhzadeh, "Deep Learning-Based Channel Estimation," in IEEE Communications Letters, vol. 23, no. 4, pp. 652-655, April 2019.

Recurrent Neural Network (RNN)

- Internal state (memory) for sequences of input
- Long Short-Term Memory (LSTM)
  - Handle exploding and vanishing gradients in traditional RNN
- Well-suited to processing time series data
  - Speech recognition
  - Machine translation
- Applications
  - CSI report (exploiting temporal correlation)
  - Source-channel coding

Long-Short Term Memory module: LSTM



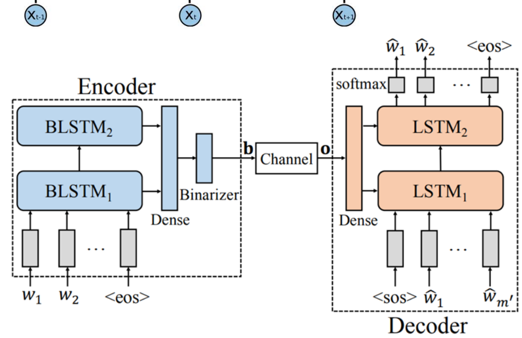
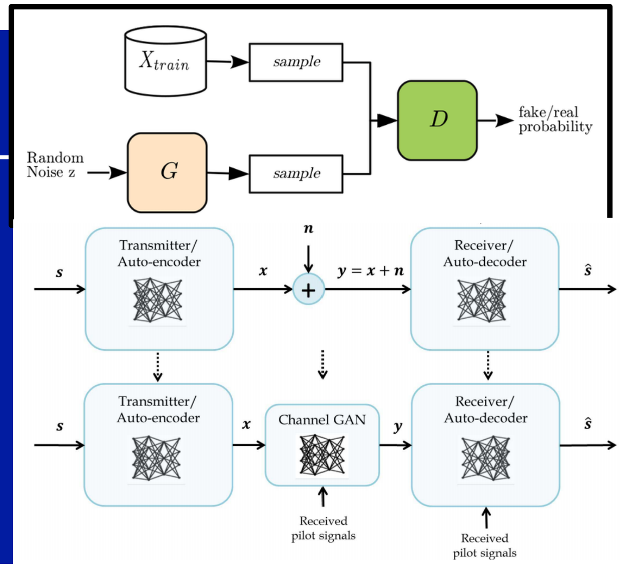


Fig. 1: The encoder-decoder architecture.

[1] Nariman Farsad, Milind Rao, and Andrea Goldsmith. "Deep learning for joint source-channel coding of text." 2018 IEEE International Conferenc[e on] Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2018.

Generative Adversarial Network (GAN)

- Generate samples indistinguishable from actual samples
- Generator
  - mimic real data distributions
- Discriminator
  - classify samples as real or fake
- Applied to mimic channel distribution



[1] Hao Ye, Le Liang, Geoffrey Ye Li, and Biing-Hwang Fred Juang. "Deep Learning based End-to-End Wireless Communication Systems with Conditional GAN as Unknown Channel." arXiv preprint arXiv:1903.02551 (2019).

Autoencoder

- Regenerate inputs
  - Dimension reduction
  - Feature learning
- Similar to source coding for recovery of transmitted data at receiver
- Applications:
  - CSI feedback compression + recovery
  - Signal detection in transceiver design

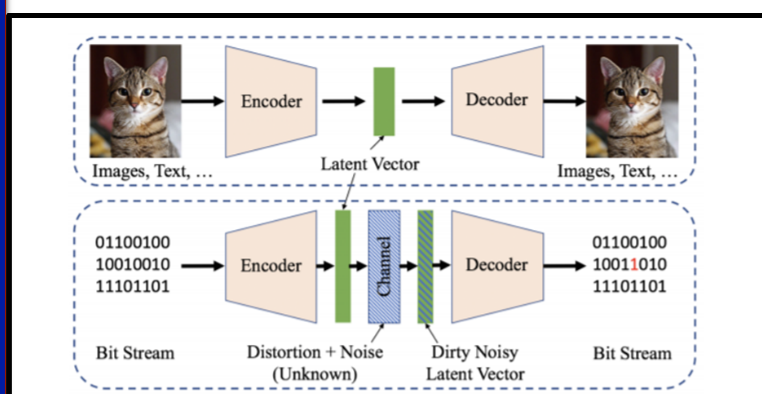
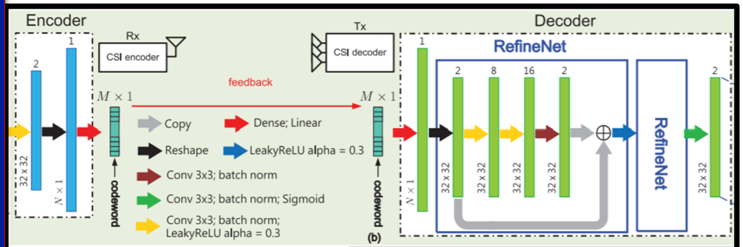


Fig. 1. General AutoEncoder (Top) v.s. Communication AutoEncoder.

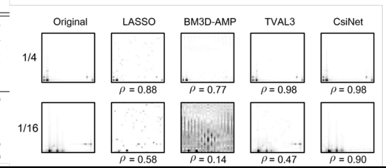
[1] Z. Liu, L. Zhang and Z. Ding, "Exploiting Bi-Directional Channel Reciprocity in Deep Learning for Low Rate Massive MIMO CSI Feedback," IEEE Wireless Communications Letters (2019).
[2] Zhongyuan Zhao. "Deep-waveform: A learned OFDM receiver based on deep complex convolutional networks." arXiv preprint arXiv:1810.07181 (2018).

Deep Learning based CSI Feedback

- CsiNet [Wen et al. 2018]
- Outperform compressive sensing (CS) methods
  - Lower delay: 50X faster
  - Higher quality
- More channel properties to exploit.



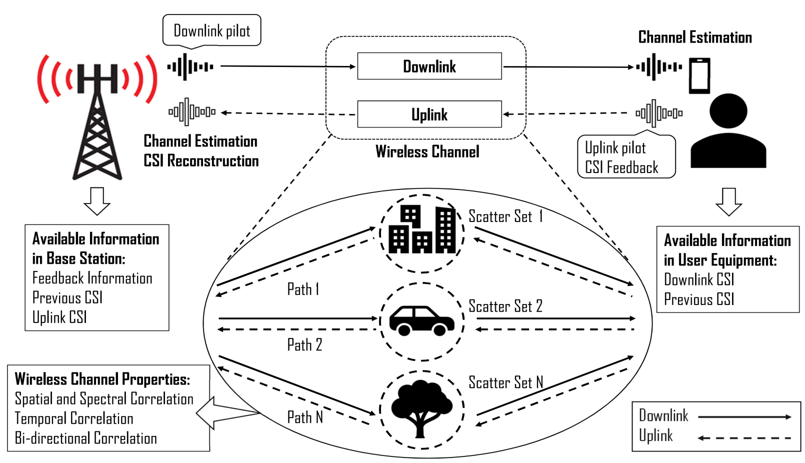
γ	Methods	Indoor NMSE	Indoor ρ	Outdoor NMSE	Outdoor ρ
1/4	LASSO	-7.59	0.91	-5.08	0.82
	BM3D-AMP	-4.33	0.80	-1.33	0.52
	TVAL3	-14.87	0.97	-6.90	0.88
	CS-CsiNet	-11.82	0.96	-6.69	0.87
	CsiNet	**-17.36**	**0.99**	**-8.75**	**0.91**
1/16	LASSO	-2.72	0.70	-1.01	0.46
	BM3D-AMP	0.26	0.16	0.55	0.11
	TVAL3	-2.61	0.66	-0.43	0.45
	CS-CsiNet	-6.09	0.87	-2.51	0.66
	CsiNet	**-8.65**	**0.93**	**-4.51**	**0.79**



[1] Chao-Kai Wen, Wan-Ting Shih, and Shi Jin. "Deep Learning for Massive MIMO CSI Feedback." IEEE Wireless Communications Letters (2018).

Deep Learning for Massive MIMO CSI Estimation

- CSI = images ?
- Complex values
- Spatial and spectral correlations
- Bi-directional CSI correlations with shared radio path
- Temporal correlations.



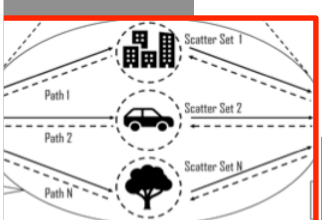
Spatial and spectral correlation

Adjacent sub-carriers and antenna elements exhibit similar propagation characteristics

Commonly exploited in compressive sensing-based CSI feedback and estimation

[1] Zhen Gao, Linglong Dai, Wei Dai, Byonghyo Shim, and Zhaocheng Wang. “Structured compressive sensing-based spatio-temporal joint channel estimation for FDD massive MIMO.” IEEE Transactions on Communications 64, no. 2 (2016): 601-617
[2] Min Soo Sim, et al. "Compressed channel feedback for correlated massive MIMO systems." Journal of Communications and Networks 18.1 (2016): 95-104.

Bi-Directional Correlation: Uplink-Downlink Radio Paths



Uplink and downlink channels experience same RF environment

CSI Reciprocity assumed in TDD

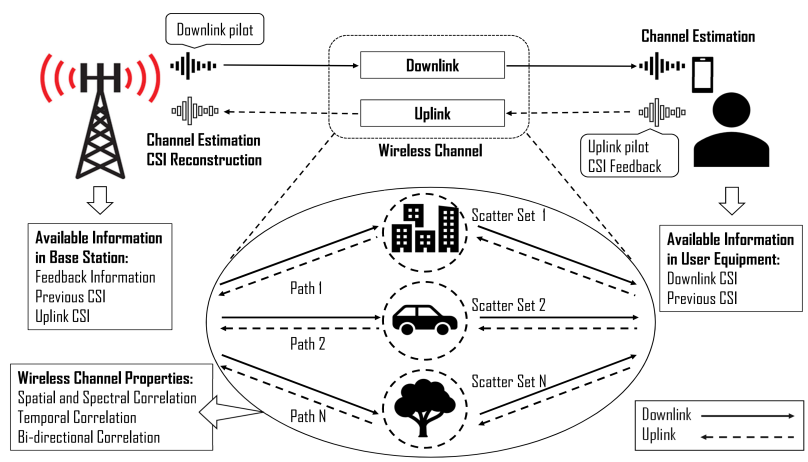
Lack of CSI reciprocity in FDD; Shared FDD channel features:

- Multipath delay profile
- Dimensionality
- Shadowing effect
- Channel covariance

[1] J. Paulraj, “Space-time processing for wireless communications,” in1997 IEEE International Conference on Acoustics, Speech, and SignalProcessing, vol. 1, April 1997, pp. 1–4 vol.1.
[2] K. Hugl, K. Kimmo, and J. Laurila, “Spatial reciprocity of uplink and downlink radio channels in fdd systems,” 2002.
[3] S. S. Szyszkowicz, H. Yanikomeroglu, and J. S. Thompson, “On the feasibility of wireless shadowing correlation models,” IEEE Trans. Veh. Technol., vol. 59, no. 9, pp. 4222–4236, Nov 2010.
[4] A. Decurninge, M. Guillaud, and D. T. M. Slock, “Channel covariance estimation in massive mimo frequency division duplex systems,” in 2015 IEEE Globecom Workshops (GC Wkshps), Dec 2015, pp. 1–6.

Shared RF Environment vs Poor CSI Reciprocity

- Physical reason?
- Similar RF paths
- Similar wavelengths
- Same time
- Same place
- Why no FDD reciprocity?



Frequeny domain vs. Delay domain

- Simple IDFT
- Benefits of delay domain:
  - Sparsity [1]
  - High entropy

[1] Zhen Gao, Linglong Dai, Wei Dai, Byonghyo Shim, and Zhaocheng Wang. “Structured compressive sensing-based spatio-temporal joint channel estimation for FDD massive MIMO.” IEEE Transactions on Communications 64, no. 2 (2016): 601-617

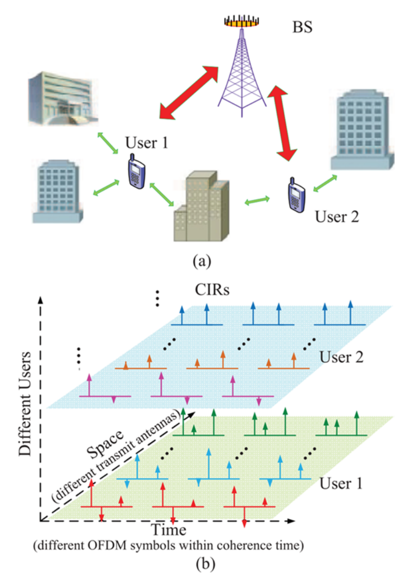
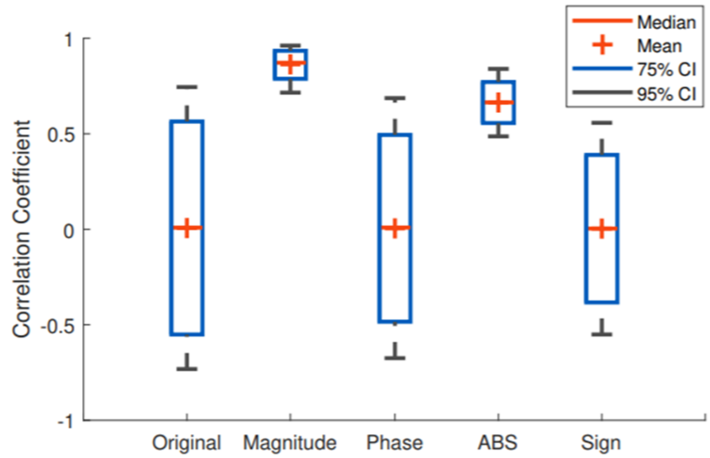


Fig. 1. Spatio-temporal common sparsity of delay-domain MIMO channels: (a) Wireless channels exhibit the sparse nature due to the limited number of

Finding Bi-Directional Correlation

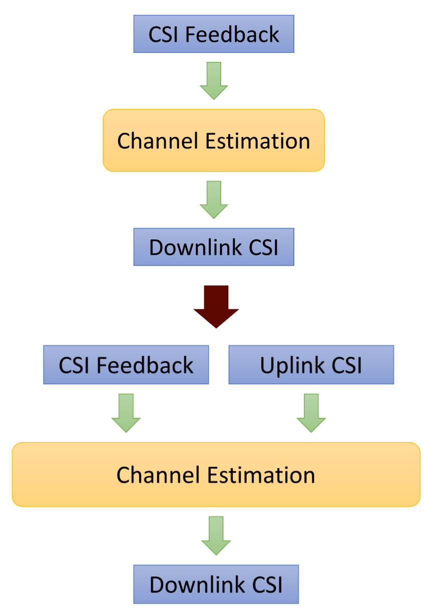
- Delay domain correlations:
  - Cartesian format ?
  - Magnitude (MAG)!!
  - Absolute value (ABS)!
  - Phase ?
  - Sign ?
- Reason? Non-coherent FDD carriers.
  - Zero phase correlation
  - Strong magnitude correlation!



DualNet

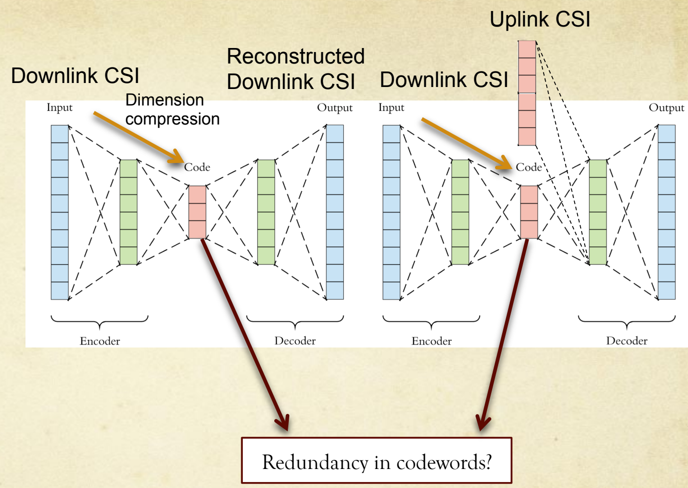
- Exploit bi-directional correlation
- DualNet-MAG
  - Correlation in magnitudes!!
- DualNet-ABS
  - Correlation in absolute values!

Zhenyu Liu, Lin Zhang and Zhi Ding, “Exploiting Bi-Directional Channel Reciprocity in Deep Learning for Low Rate Massive MIMO CSI Feedback," IEEE Wireless Communications Letters (2019).



Deep Learning based CSI Feedback

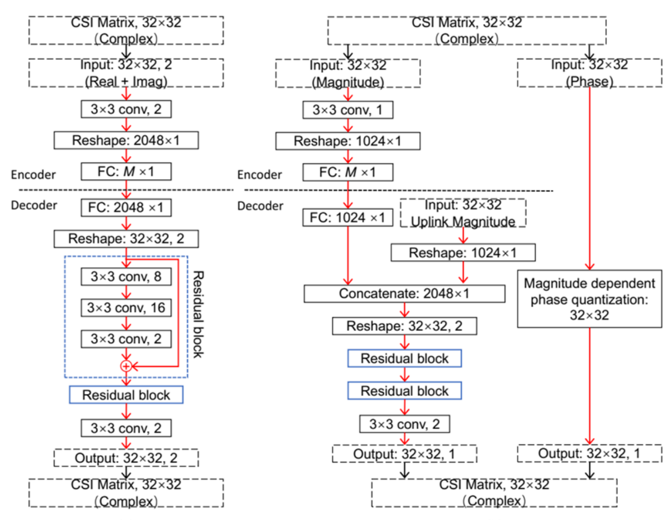
- From CsiNet to ....
- DualNet
- Outperform compressive sensing (CS)-based methods
  - Lower delay: 50 times faster
  - Higher reconstruction quality



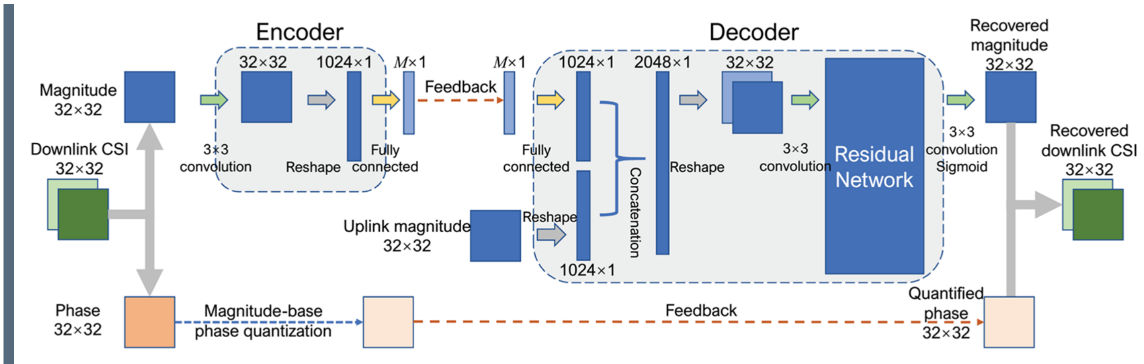
[1] Chao-Kai Wen, Wan-Ting Shih, and Shi Jin. "Deep Learning for Massive MIMO CSI Feedback." IEEE Wireless Communications Letters (2018).
[2] Z. Liu, L. Zhang and Z. Ding, "Exploiting Bi-Directional Channel Reciprocity in Deep Learning for Low Rate Massive MIMO CSI Feedback," IEEE Wireless Communications Letters (2019).

Deep Learning based CSI Feedback

- From CsiNet to ...
- DualNet
- Outperform compressive sensing (CS)-based methods
  - More efficient model

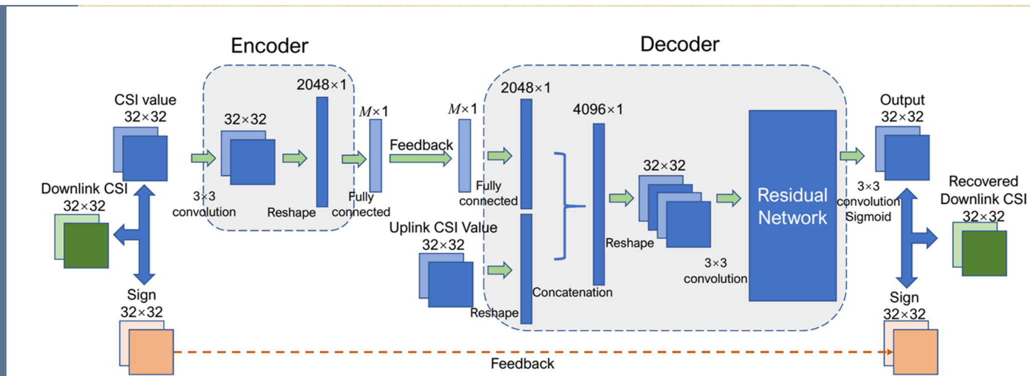


[1] Chao-Kai Wen, Wan-Ting Shih, and Shi Jin. "Deep Learning for Massive MIMO CSI Feedback." IEEE Wireless Communications Letters (2018).
[2] Z. Liu, L. Zhang and Z. Ding, "Exploiting Bi-Directional Channel Reciprocity in Deep Learning for Low Rate Massive MIMO CSI Feedback," IEEE Wireless Communications Letters (2019).



DualNet-MAG

- Magnitudes and phases of CSI separately encoded
- Magnitude dependent phase quantization



DualNet-ABS

- Sizes and signs of real/imaginary CSI separated
- Signs save more space than phases

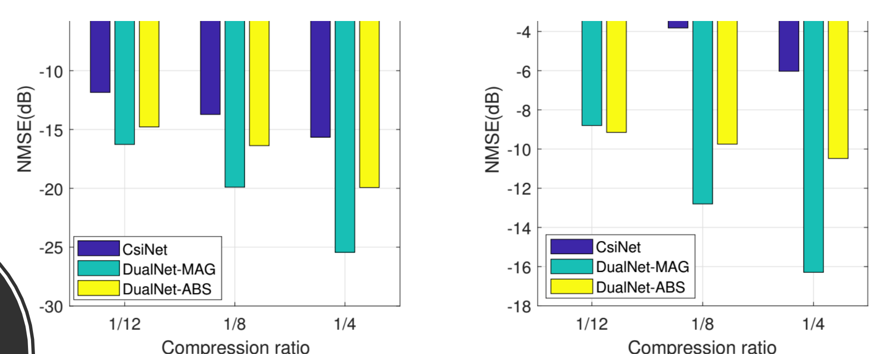
MIMO Scenarios

- COST 2100 model
  - Frequency
    - Indoor: Downlink: 5.3 GHz Uplink: 5.1 GHz
    - Outdoor: Downlink: 300MHz Uplink: 260MHz
  - Bandwidth: 20 MHz
  - Subcarrier: 1024
  - Antennas: 32

[1] L. Liu, J. Poutanen, F. Quitin, K. Haneda, F. Tufvesson, P. De Doncker, P. Vainikainen and C. Oestges, “The COST 2100 MIMO channel model,” IEEE Wireless Commun., vol 19, issue 6, pp 92-99, Dec. 2012.

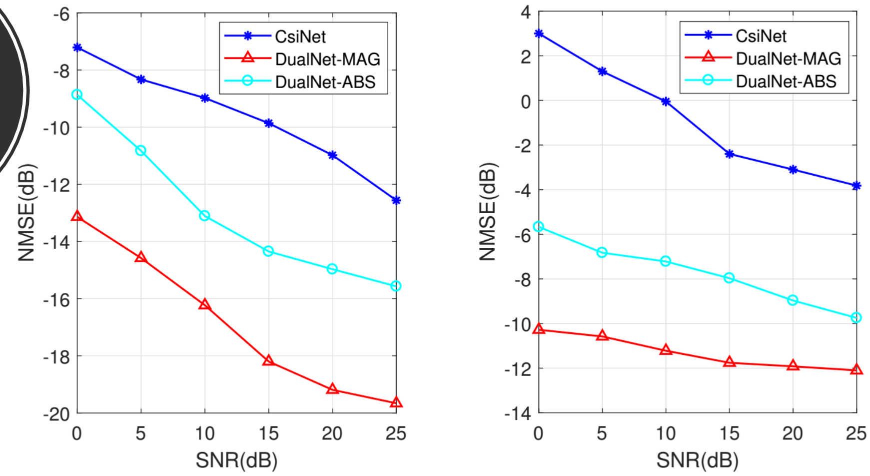
Performance Evaluation

Bi-directional correlation brings obvious improvement



(a) Indoor (b) Outdoor

Performance Evaluation



(a) Indoor (b) Outdoor

Fig. 4: CSI recovery comparison at different CSI error levels

CsiNet, DualNet-MAG and DualNet-ABS:

Similar complexity

Table 1: Comparisons among CsiNet, DualNet-MAG and DualNet-ABS when CR = 1/8

Architecture	Layers	Parameters	Training time/epoch	Testing time/sample
CsiNet	28	1,054,624	11s	0.000094s
DualNet-MAG	30	266,993	14s	0.000086s
DualNet-ABS	30	802,008	24s	0.000101s

Complexity Comparison

Phase Quantization

- Poor performance when using deep network model
- Uniform phase quantization (UQ)
  - Unnecessarily fine quantization
- Magnitude dependent phase quantization (MDPQ)

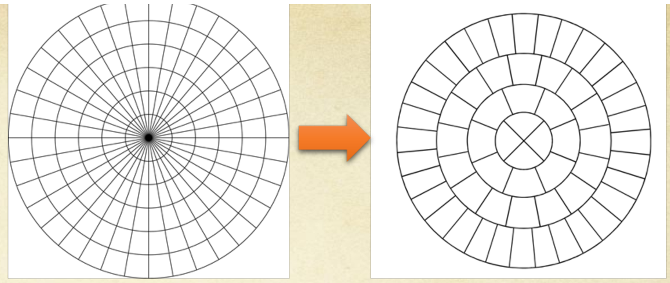
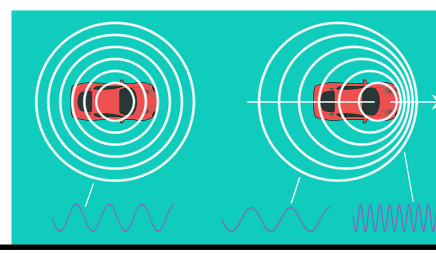


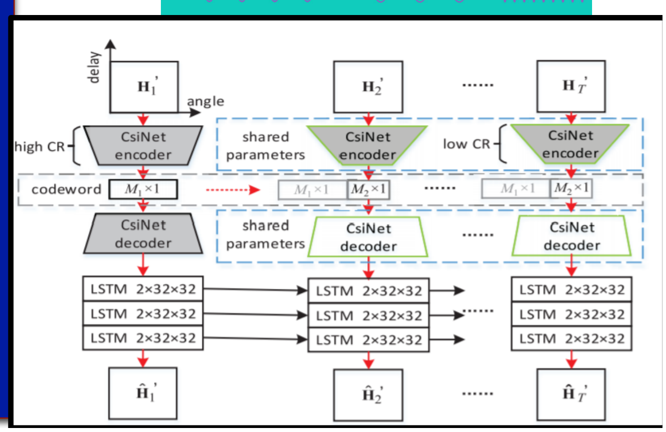
TABLE I: NMSE comparison between UQ and MDPQ

Scenario	UQ			MDPQ
	5	6	7	4
Indoor	-19.0	-19.9	-20.1	-19.9
Outdoor	-12.6	-12.8	-12.9	-12.8

LSTM for Exploiting Temporal Correlation

- Doppler spread
- Coherence time
- Perfectly match capabilities of RNN and LSTM
- CsiNet-LSTM

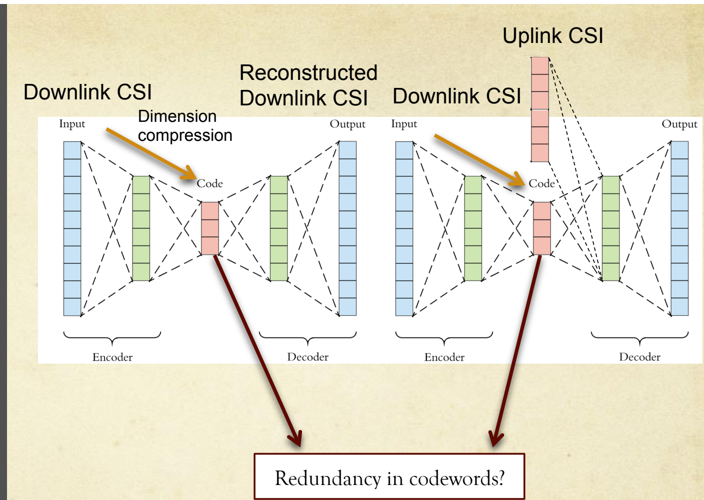




[1] T. Wang, C. Wen, S. Jin and G. Y. Li, "Deep Learning-Based CSI Feedback Approach for Time-Varying Massive MIMO Channels," in IEEE Wireless Communications Letters, vol. 8, no. 2, pp. 416-419, April 2019.

Deep Learning based CSI Feedback

- Information-Theoretic View: Entropy compression.



[1] Chao-Kai Wen, Wan-Ting Shih, and Shi Jin. "Deep Learning for Massive MIMO CSI Feedback." IEEE Wireless Communications Letters (2018).
[2] Z. Liu, L. Zhang and Z. Ding, "Exploiting Bi-Directional Channel Reciprocity in Deep Learning for Low Rate Massive MIMO CSI Feedback," IEEE Wireless Communications Letters (2019).

Much Simpler DiffNet (in Preparation)

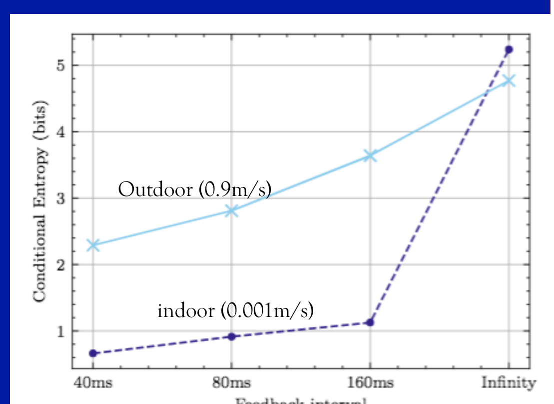
- Exploiting temporal correlation
- Markov Model

$$H(\mathbf{H}_t|\mathbf{H}_{t-1},\ldots,\mathbf{H}_1) = H(\mathbf{H}_t|\mathbf{H}_{t-1})$$

- Leveraging CSI at *t-1* to reduce feedback:

$$H(\mathbf{H}_t|\mathbf{H}_{t-1}) \leq H(\mathbf{H}_t)$$

- Substantially reduced CSI feedback

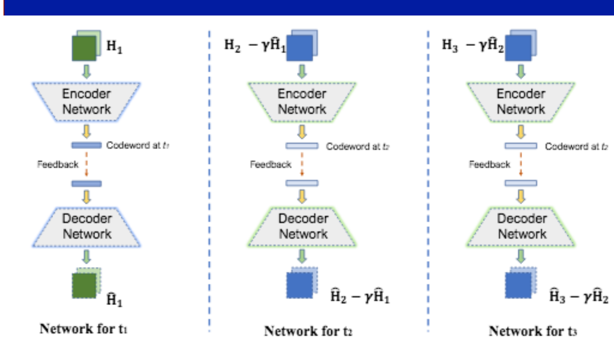


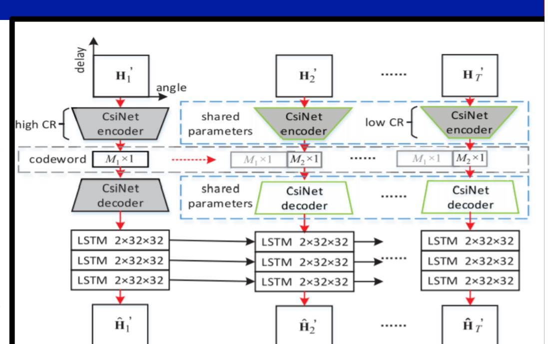
DiffNet (2)

- MMSE Prediction $\hat{H}_t = E\{H_t \mid H_{t-1}\} = \Psi(H_{t-1})$
- Simplified Markov model $\mathbf{H}_t = \gamma \mathbf{H}_{t-1} + \mathbf{V}_t$
- Encoder: Autoencoder to obtain feedback $\mathbf{s}_t \;=\; f_{e,t}(\mathbf{H}_t - \gamma \hat{\mathbf{H}}_{t-1})$ based on previous CSI estimate
- Decoder: Reconstruct CSI $\hat{\mathbf{H}}_t \;=\; f_{d,t}(\mathbf{s}_t) + \gamma \hat{\mathbf{H}}_{t-1}$ using feedback St.

DiffNet vs LSTM

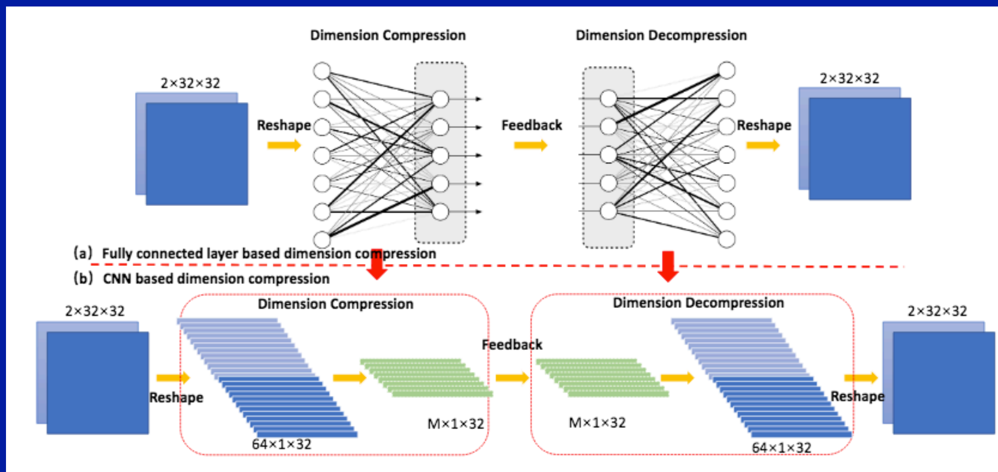
- Exploiting CSI correlation and Conditional entropy directly
- Proposal of simple Markov model => structured RNN





CNN to the Rescue in DiffNet

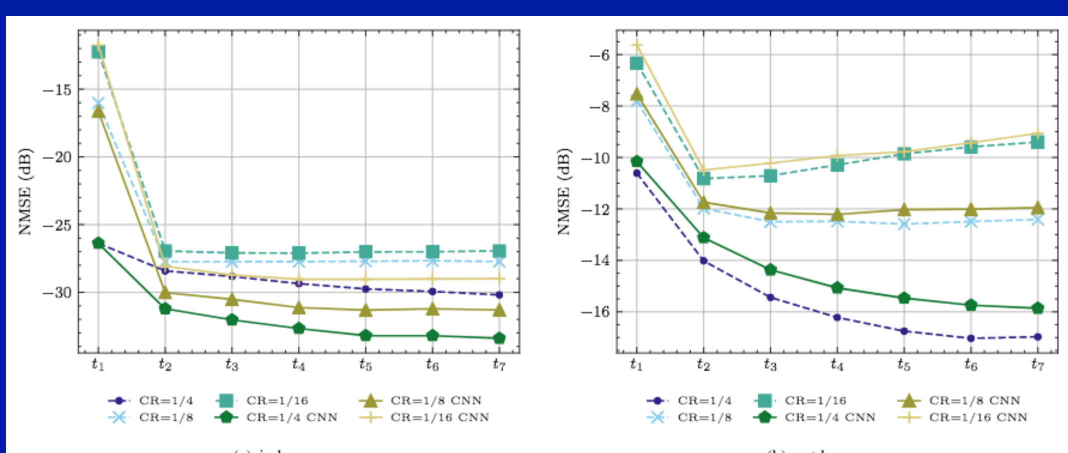
- CNN substantially simplifies the DiffNet encoder/decoder



-

CNN to the Rescue in DiffNet

- ... at little performance loss



-

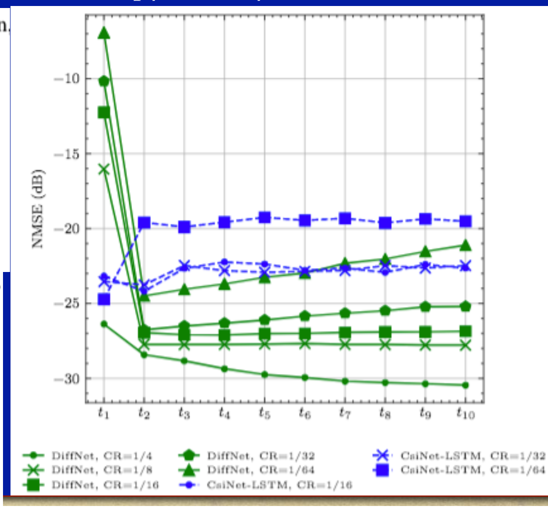
Indoor Channel Performance

- Exploiting CSI correlation and Conditional entropy directly
- TABLE III: Number of parameters comparison. M: million. K: thousand.

	CsiNet-LSTM	DiffNet	DiffNet-CNN
CR=$\frac{1}{4}$	132.7 M	2.1 M	34.9 K
CR=$\frac{1}{8}$	123.2 M	1.1 M	27.8 K
CR=$\frac{1}{16}$	118.5 M	542.9 K	24.2 K
CR=$\frac{1}{32}$	116.1 M	280.7 K	22.4 K
CR=$\frac{1}{64}$	115.0 M	149.6 K	21.5 K

TABLE IV: Number of FLOPs comparison. M: million.

	CsiNet-LSTM	DiffNet	DiffNet-CNN
CR=$\frac{1}{4}$	412.9 M	44.5 M	41.2 M
CR=$\frac{1}{8}$	410.8 M	42.4 M	40.7 M
CR=$\frac{1}{16}$	409.8 M	41.3 M	40.5 M
CR=$\frac{1}{32}$	409.2 M	40.8 M	40.4 M
CR=$\frac{1}{64}$	409.0 M	40.5 M	40.3 M



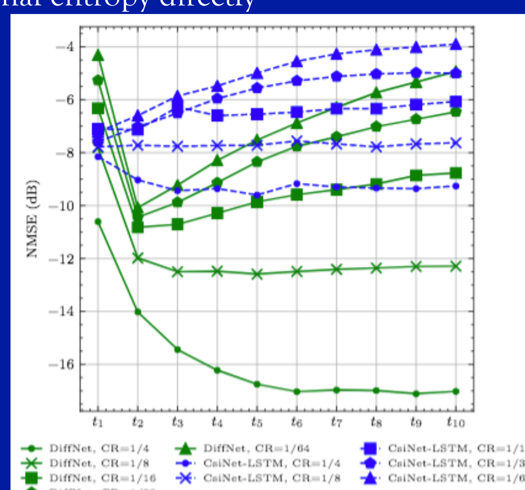
Outdoor Channel Performance

- Exploiting CSI correlation and Conditional entropy directly
- TABLE III: Number of parameters comparison. M: million. K: thousand.

	CsiNet-LSTM	DiffNet	DiffNet-CNN
CR=$\frac{1}{4}$	132.7 M	2.1 M	34.9 K
CR=$\frac{1}{8}$	123.2 M	1.1 M	27.8 K
CR=$\frac{1}{16}$	118.5 M	542.9 K	24.2 K
CR=$\frac{1}{32}$	116.1 M	280.7 K	22.4 K
CR=$\frac{1}{64}$	115.0 M	149.6 K	21.5 K

TABLE IV: Number of FLOPs comparison. M: million.

	CsiNet-LSTM	DiffNet	DiffNet-CNN
CR=$\frac{1}{4}$	412.9 M	44.5 M	41.2 M
CR=$\frac{1}{8}$	410.8 M	42.4 M	40.7 M
CR=$\frac{1}{16}$	409.8 M	41.3 M	40.5 M
CR=$\frac{1}{32}$	409.2 M	40.8 M	40.4 M
CR=$\frac{1}{64}$	409.0 M	40.5 M	40.3 M



Open Issues

- DL under different system settings
- Specialized DL models combined with channel features
- Low complexity & distributed DL in wireless nodes
- Tradeoff between performance and training efficiency
- Transfer learning based approaches

DL Under Different System settings

- Wireless system parameters affect CSI correlation and DL performance
  - UL/DL Band-gap
  - Bandwidth
  - Carrier frequency

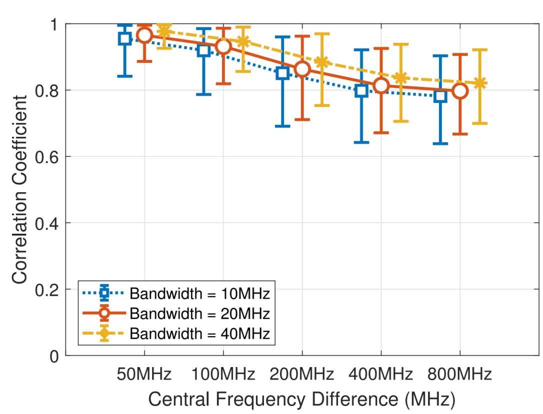
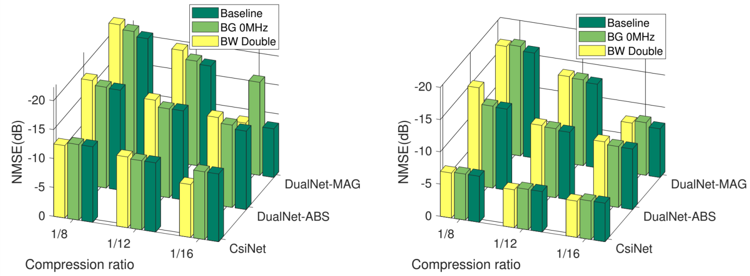


Fig. 3: Influence of band gap and bandwidth on channel magnitude correlation within 95% CI.

DL under different system settings

- Stronger bi-directional correlation =>
  better DualNet performance



(a) Indoor (b) Outdoor

Fig. 5: CSI feedback performance in the influence of band gap (BG) and bandwidth (BW).

Conclusions

A good match-making is based on the characteristics of both parties!

- [1] Zhenyu Liu, Lin Zhang and Zhi Ding, "Exploiting Bi-Directional Channel Reciprocity in Deep Learning for Low Rate Massive MIMO CSI Feedback," *IEEE Wireless Communications Letters* (2019).
- [2] Z. Liu, L. Zhang, and Z. Ding, “An Efficient Deep Learning Framework for Low Rate massive MIMO CSI Reporting,” IEEE Transactions on Communications, 2020.
- [3] Z. Liu, M. del Rosario, X. Liang, L. Zhang, and Z. Ding, “Spherical Normalization for Learned Compressive Feedback in Massive MIMO CSI Acquisition,” 2020 IEEE International Conference on Communications, Machine Learning for Communications Workshop, May 2020.
- [4] Zhenyu Liu, Lin Zhang and Zhi Ding, "Massive MIMO Meets Deep Learning: Clearing the CSI Barrier for Future Wireless Communications," *submitted to IEEE Communications Magazine,* arXiv:1912.10573, 2019

Codes to SHARE

Open Source Code for DualNet and U2D Networks

- DualNet_abs.py
- DualNet_mag.py
- Introduction to U2D-ORG U2D-ABS and U2D-MAG.pdf
- NMSE_performance.m
- README.md
- u2d_abs.py
- u2d_mag.py
- u2d_org.py

https://ieee-collabratec.ieee.org/app/groups/6247/Bi-directional-Channel-Code-Sharing/activities

Thank you

Special thanks to PhD students Zhenyu Liu and Mason del Rosario