Full Duplex Communications for the Next Generation Wireless Networks

Prof. Octavia A. Dobre

Electrical and Computer Engineering
Faculty of Engineering and Applied Science
Memorial University, Canada

Acknowledgments

MEMORIAL UNIVERSITY

I would like to thank Dr. Khaled M. Fouad Esayed and Dr. Hossam A. Abdelfattah for the invitation to give a talk at

This work has been supported in part by the Huawei Technologies Canada.

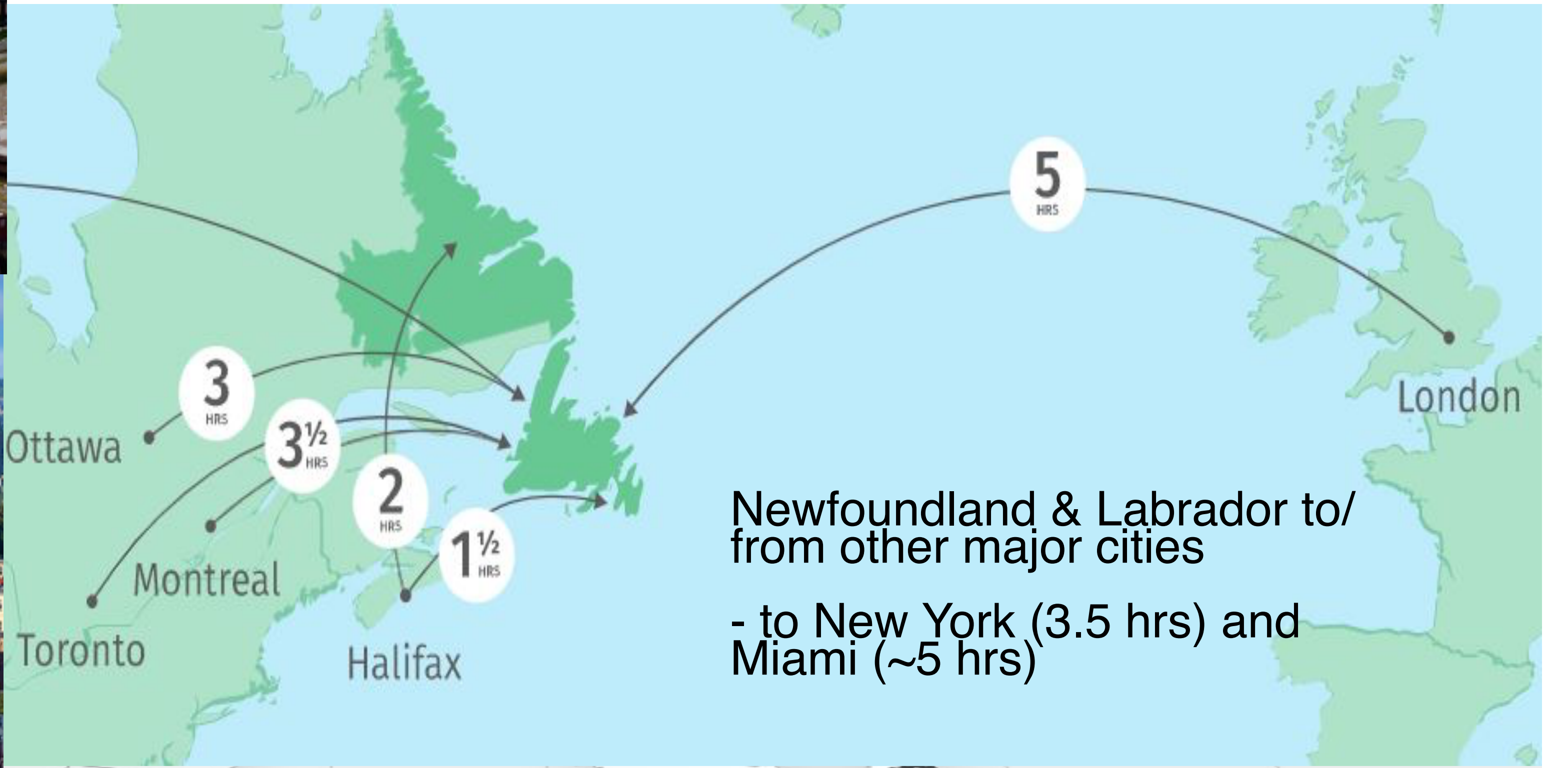
- M. Elsayed, A. A. El-Banna, O. A. Dobre, W. Shiu, and P. Wang, "Machine learning-based self-interference cancellation for full-duplex radio: Approaches, open challenges, and future research directions," (Invited paper), *IEEE Open Journal of Vehicular Technology*, Apr. 2023.

- M. Elsayed, A. A. El-Banna, O. A. Dobre, W. Shiu, and P. Wang, "Hybrid-layers neural network architectures for modeling the self-interference in full-duplex systems," *IEEE Transactions on Vehicular Technology*, vol. 71, issue 6, pp. 6291-6307, June 2022.

- M. Elsayed, A. A. El-Banna, O. A. Dobre, W. Shiu, and P. Wang, "Full-duplex self-interference cancellation using dual-neurons neural networks," *IEEE Communications Letters*, vol. 26, issue 3, pp. 557-561, Mar. 2022.

- M. Elsayed, A. A. El-Banna, O. A. Dobre, W. Shiu, and P. Wang, "Low complexity neural network structures for self-interference cancellation in full-duplex radio," *IEEE Communications Letters*, vol. 25, issue 1, pp.181-185, Jan. 2021.

Where we are located



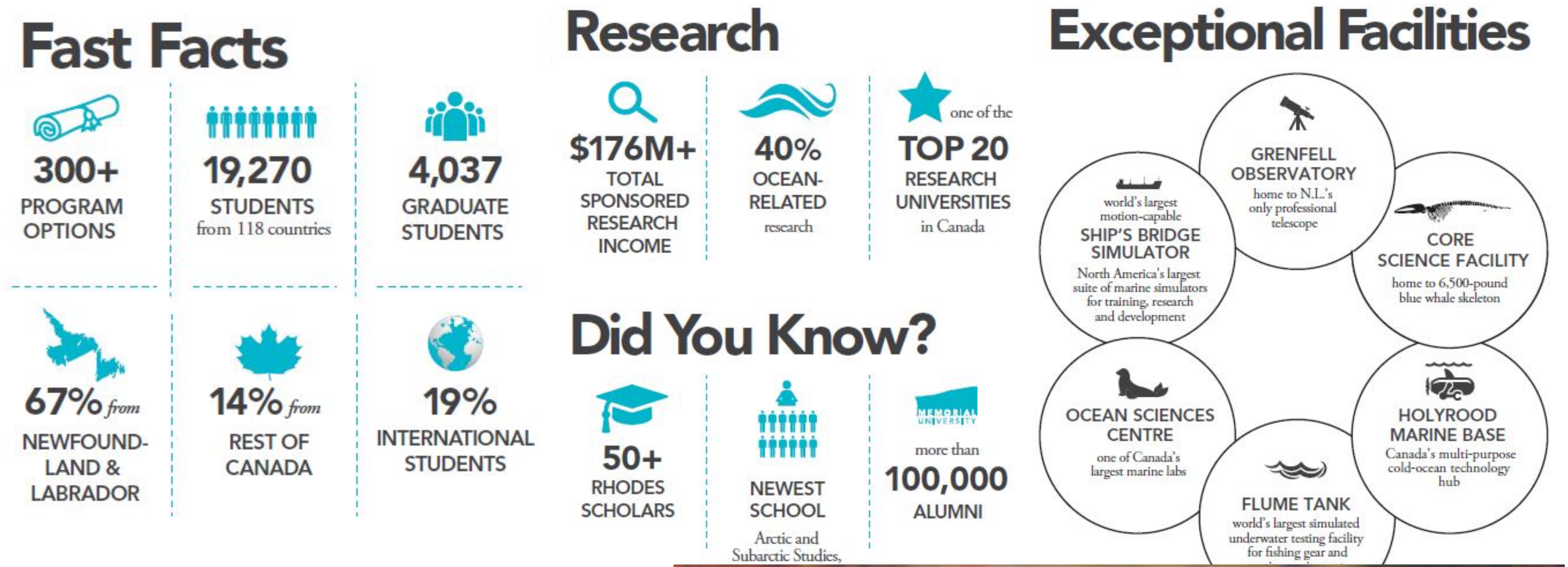
Newfoundland & Labrador to/ from other major cities

- to New York (3.5 hrs) and Miami (~5 hrs)

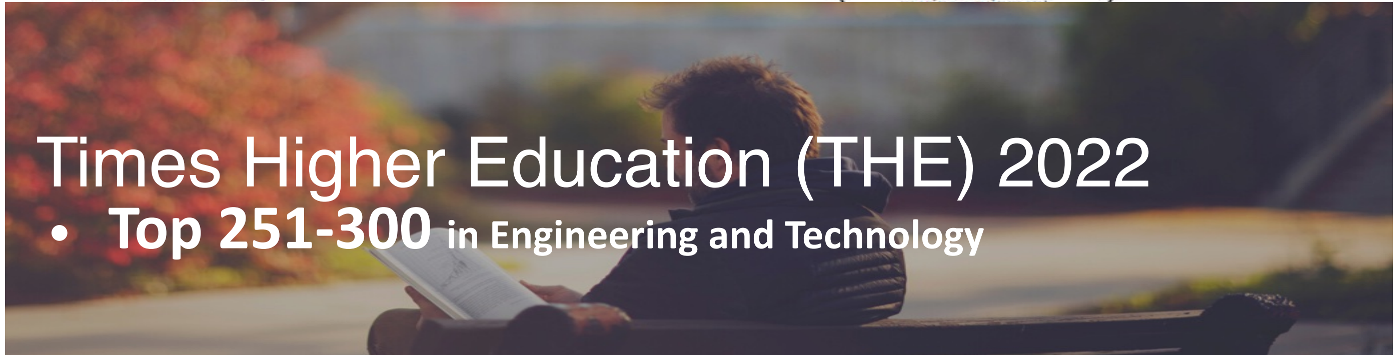
& Labrador

MEMORIAL UNIVERSITY

Memorial University – the largest university in Atlantic Canada



Faculty of Engineering and Applied Science



Times Higher Education (THE) 2022

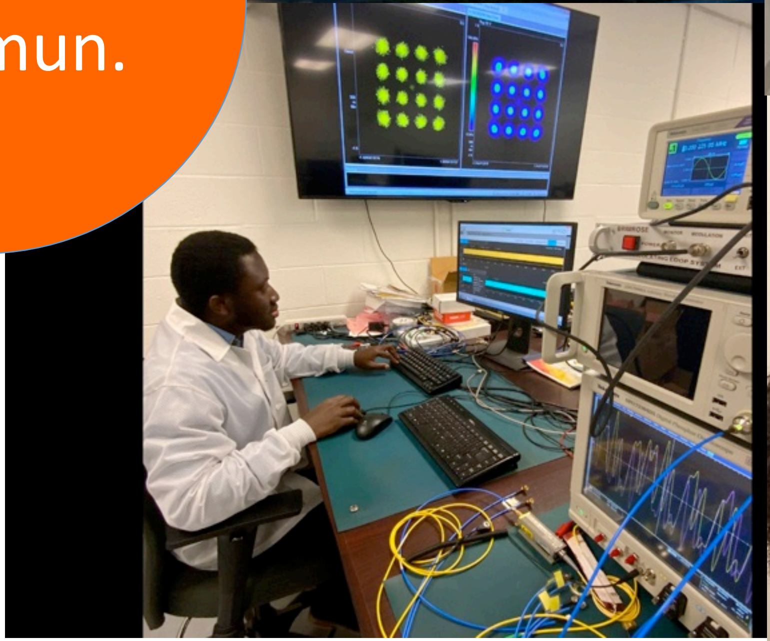
- **Top 251-300** in Engineering and Technology

Research Overview

Wireless Commun.

Optical Commun.

Underwater Commun.



[Memorial's Advanced Research Laboratory on Communications]()

Research Overview

MEMORIAL UNIVERSITY

Funding ~16 million dollars

Sources: NSERC, MITACS, CFI, ACOA, InnovateNL/RDC,
DRDC, CRC, DoD, Statoil/Equinor, Altera/Intel Canada, Huawei Tech. Canada, Agile Tech.
EION, Allen Vanguard, ThinkRF, DTA Systems, Keithley/Agilent

Collaborations

Memorial: +10 faculty members – different departments in Engineering, Computer Science, Mathematics & Statistics

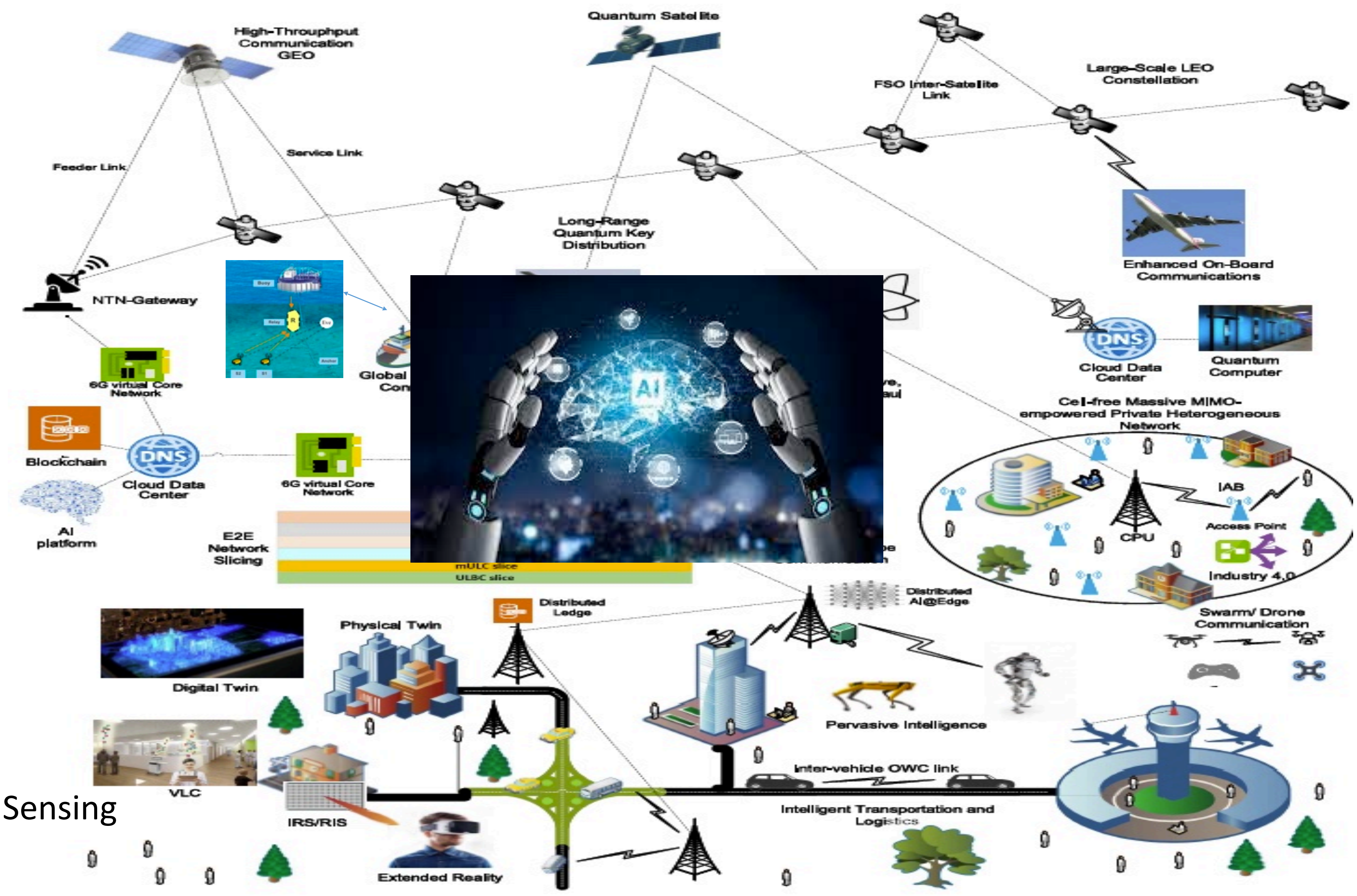
Canada: UBC, University of Toronto, Dalhousie University

International: 15 institutions in 10 countries

Research Topics

- Artificial Intelligence for Communications
- Wireless Communications:
  - Technologies for Beyond 5G – 6G Wireless Networks: ISAC, RIS/IRS, NTN, FD
  - Resource Allocation Designs in Wireless Networks
  - Blind Signal Identification
- Optical Communications: Parameter Estimation and Non-linearity Compensation in Long-Haul Optical Networks
- Underwater Communications: Channel Estimation, FD, NOMA

The Hyper-Connected Future World: NextG Networks



Sensing

Key Areas

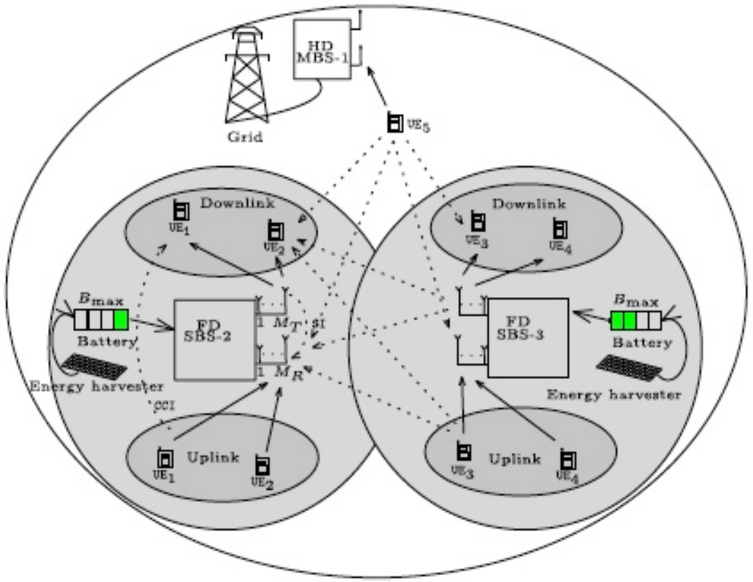
- Intelligent systems
- Digital twin
- THz devices & communications
- Intelligent reflective surface
- Non-terrestrial networks & Internetworking networks
- Integrated sensing and comm.
- Holographic-type communications
- Quantum comms & computing
- Security and privacy

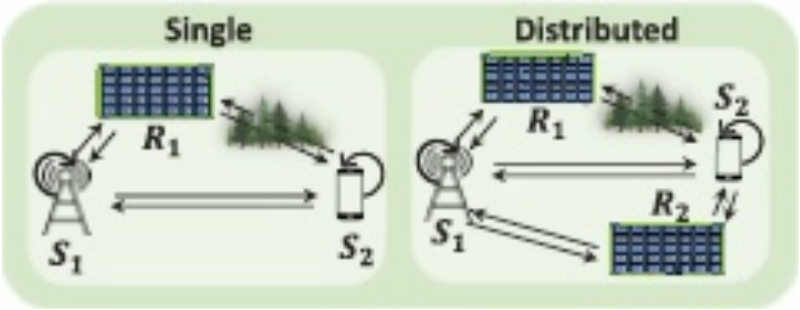
W. Jiang et al., “The road towards 6G: A comprehensive survey,” *IEEE Open Journal of Communications Society*, pp. 334-366, Feb. 2021.

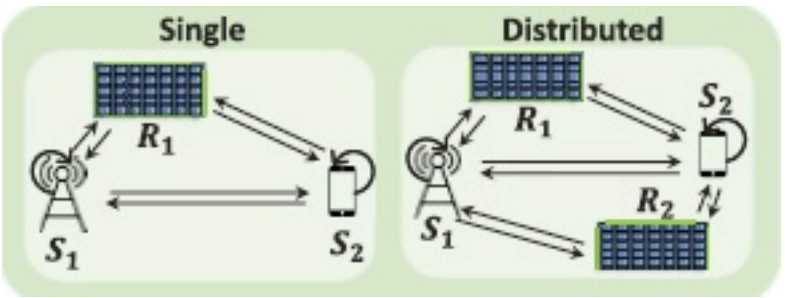
Outline

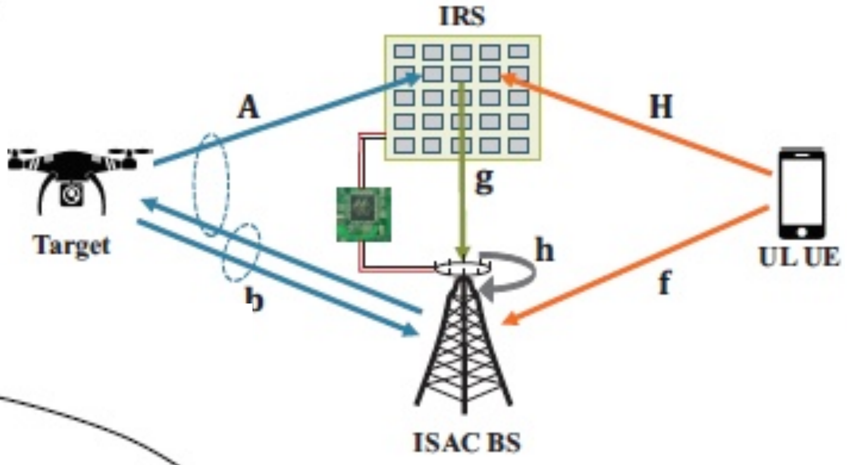
- Full-duplex Communications
- Self-interference Cancellation (SIC) in Full-duplex Transceivers
    - Full-duplex transceiver model
    - Neural network (NN)-based SIC
    - Support vector regressor (SVR)-based SIC
    - Achieved results & comparisons (with other methods)
- Summary, Conclusions, and Future Work

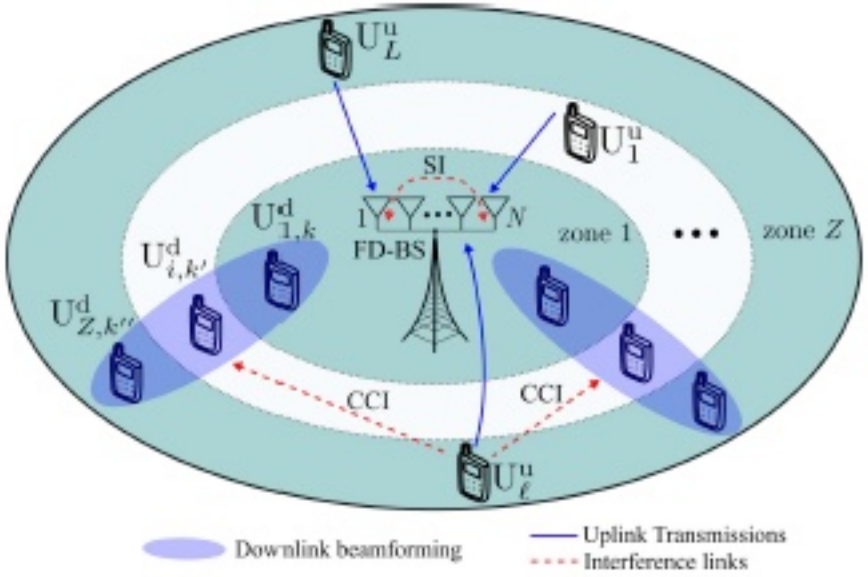
Full-Duplex Communications

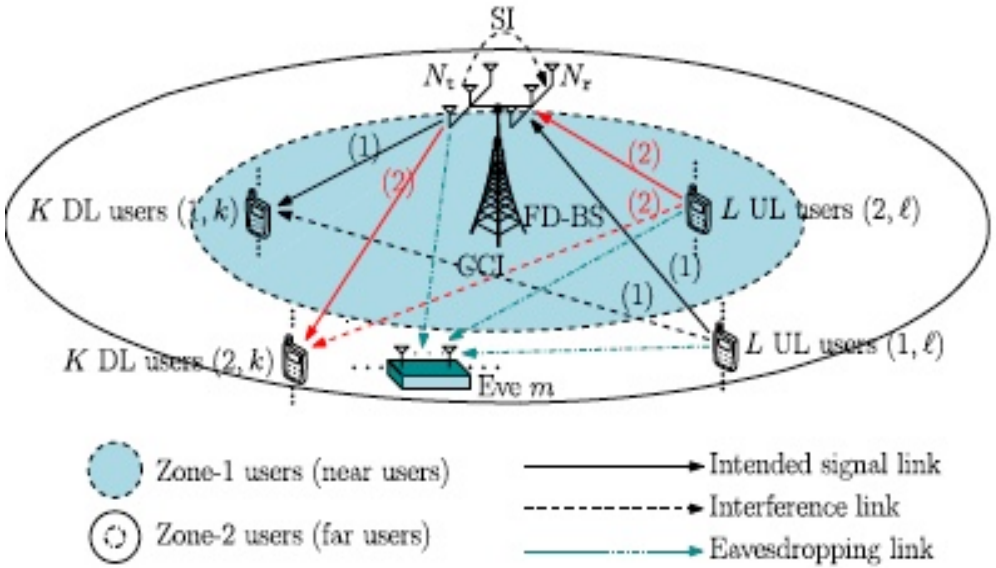






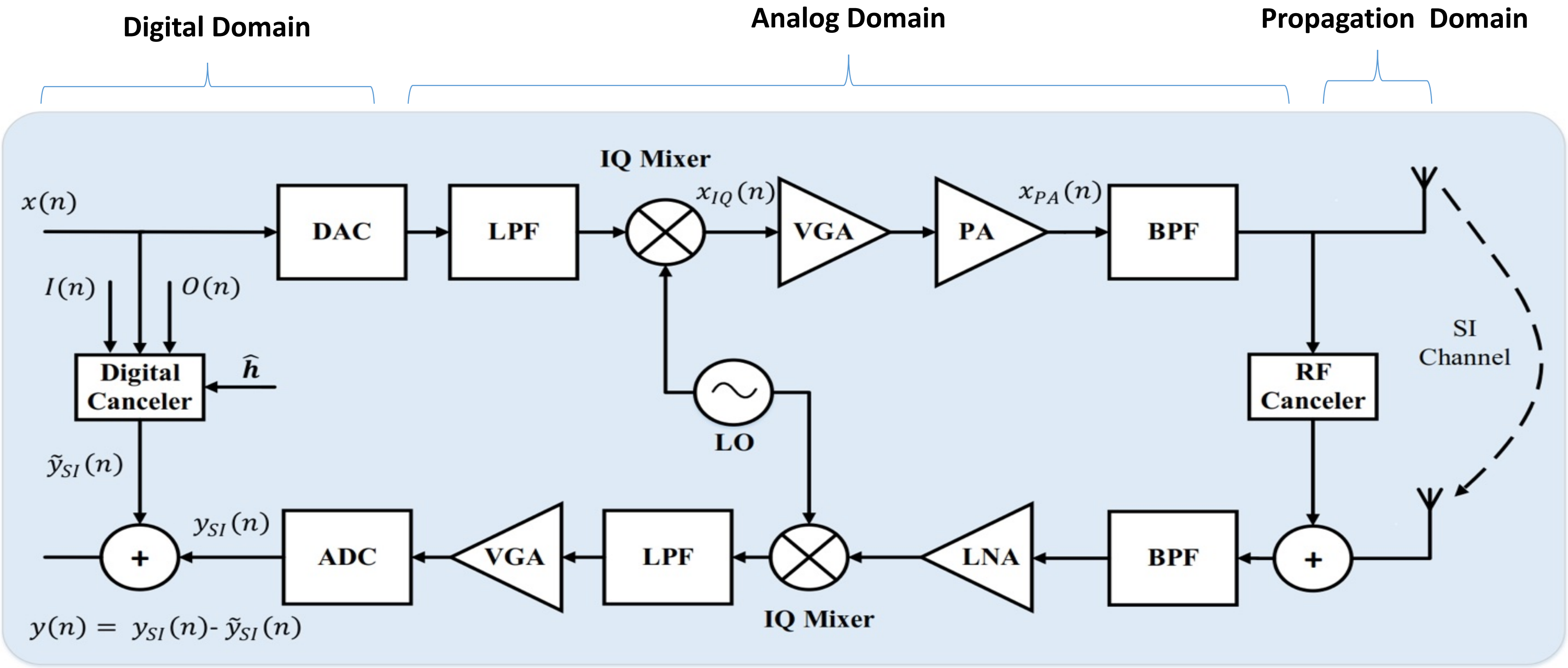






- H. V. Nguyen et al., “Joint power control and user association for NOMA-based full-duplex systems,” *IEEE Transactions on Communications*, Nov. 2019.
- A. Faisal et al., “Deep reinforcement learning for RIS-assisted FD systems: Single or distributed RIS?,” *IEEE Communications Letters*, July 2022.
- Y. Liu et al., “Deep-learning channel estimation for IRS-assisted integrated sensing and communication system,” *IEEE Transactions on Vehicular Technology*, Dec. 2022.

Self-Interference: Full-Duplex Transceiver Model



FD transceiver model with two stage cancellation techniques.

DAC: digital-to-analog converter; **LPF:** low pass filter; **VGA:** variable gain amplifier;
PA: power amplifier; **BPF:** band pass filter; **LNA;** low noise amplifier;
ADC: analog-to-digital converter; **LO:** local oscillator.

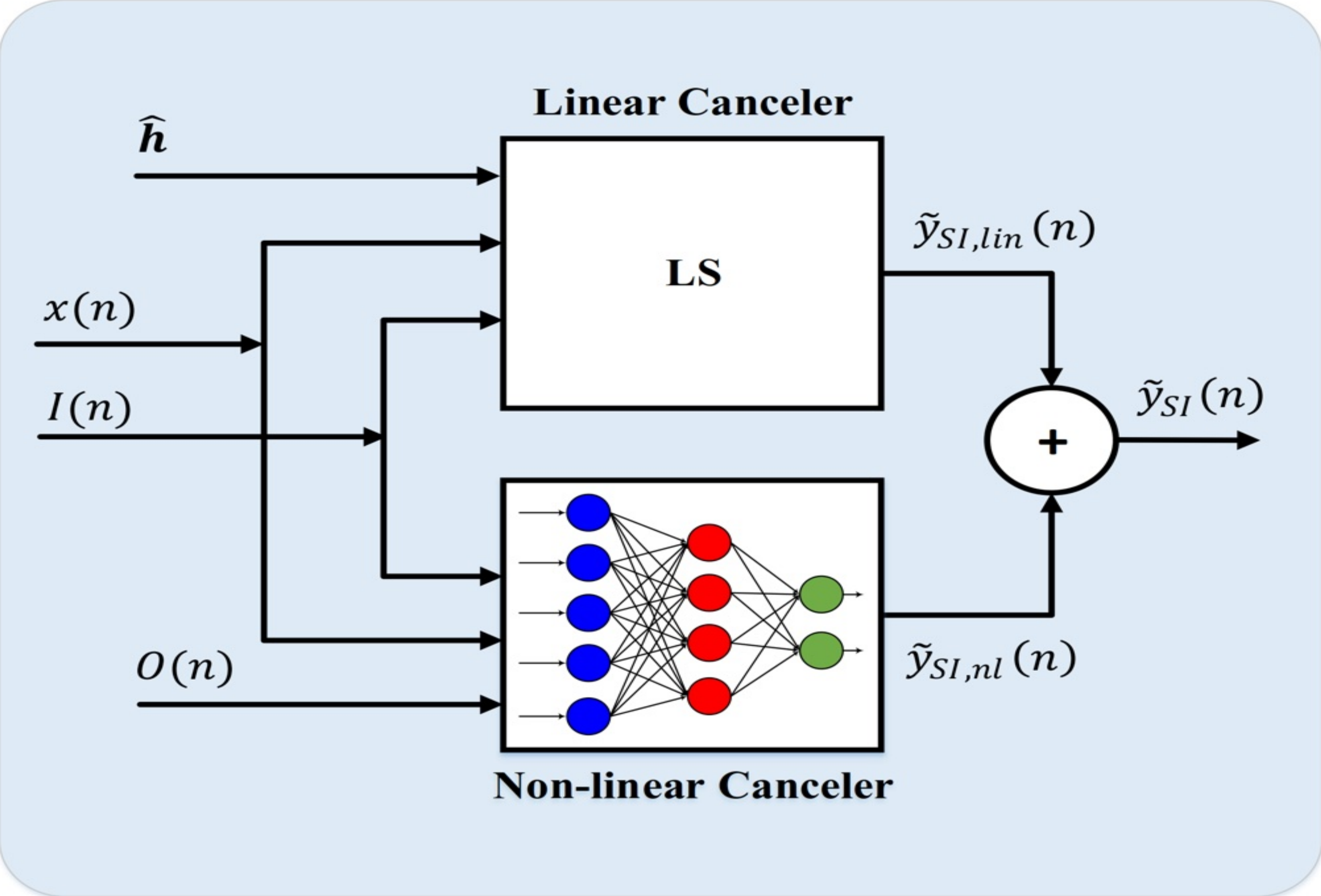
Non-linearity Sources

- PA and LNA non-idealities
- IQ imbalance
- Phase noise
- Quantization noise

- Y. Kurzo, A. T. Kristensen, A. Burg, and A. Balatsoukas-Stimming, “Hardware Implementation of Neural Self-Interference Cancellation,” *IEEE J. Emerg. Sel. Topics Circuits Syst.*, Jun. 2020.

Full-Duplex Transceiver Model

➢ Digital Canceler



Digital canceler.

- Total cancellation: linear plus non-linear cancellation

- $I(n) = \{x(n-1), x(n-2), ..., x(n-M_i+1)\}$
- $O(n) = \{z(n-1), z(n-2), ..., z(n-M_o)\}$
- $z(n) = y_{SI}(n) - \tilde{y}_{SI,lin}(n)$

- The modeled SI signal is decomposed into:
  - **Linear part:** estimated using the conventional linear cancellation which is based on the least-square (LS) channel estimation
  - **Non-linear part:** approximated using machine learning, e.g., NN, SVR

Existing NN-based SIC methods

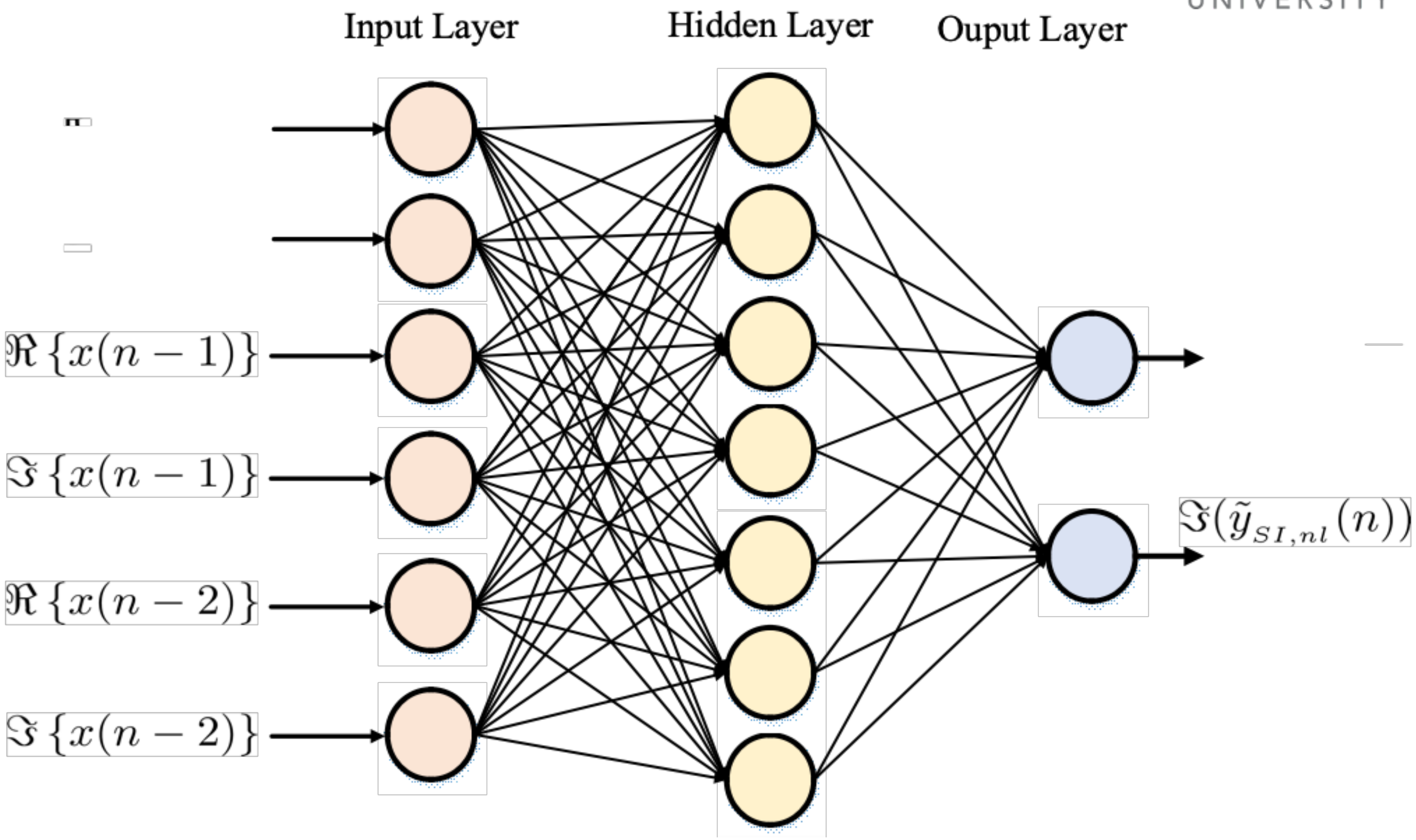
- Real-valued time delay neural network (RV-TDNN)
- Recurrent neural network (RNN)
- Complex-valued time delay neural network (CV-TDNN)

Recent NN-based SIC methods

- **Real-valued TDNN (RV-TDNN)**

- Previously investigated for behavioral modelling of Pas

- Also investigated in SIC problem
- It can match the performance of the polynomial non-linear canceler with lower computational complexity

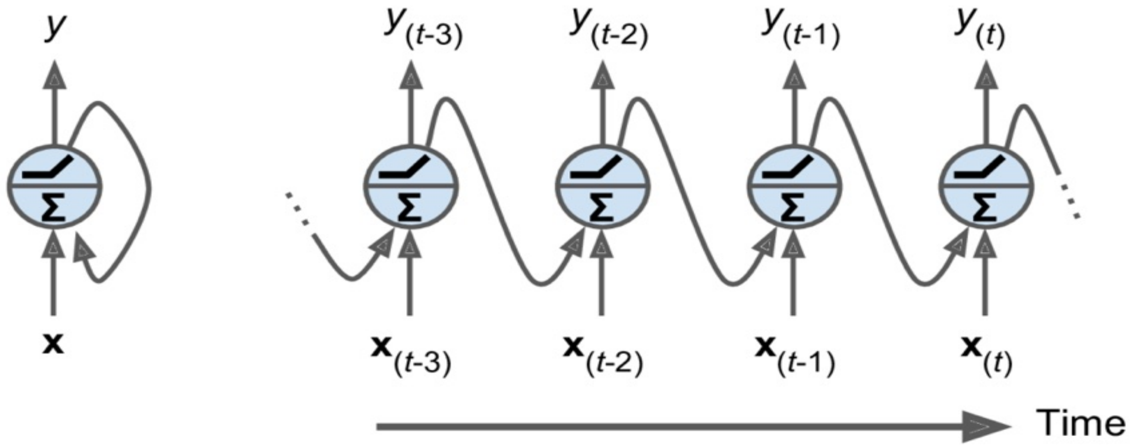


RV- TDNN architecture.

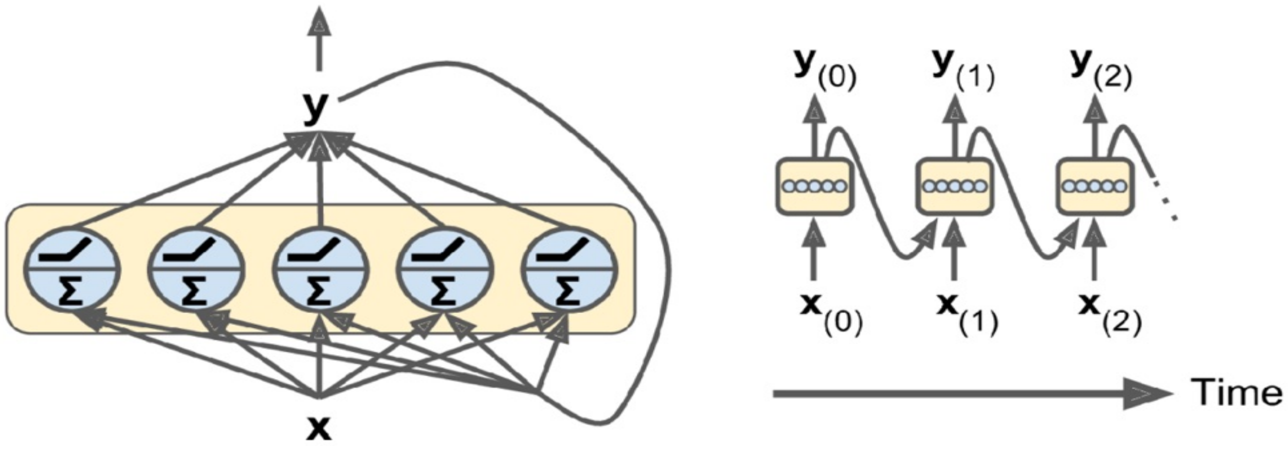
- A. Balatsoukas-Stimming, "Non-linear digital self-interference cancellation for in-band full-duplex radios using neural networks," in *Proc. IEEE Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2018.

Recent NN-based SIC methods

- **Recurrent NN (RNN)**
  - Connections pointing backward
  - Requires high training complexity, which makes it unpopular for real-time deployment



(a) A recurrent neuron (left) unrolled through time (right)



(b) A layer of recurrent neurons (left) unrolled through time (right)

RNN architecture.

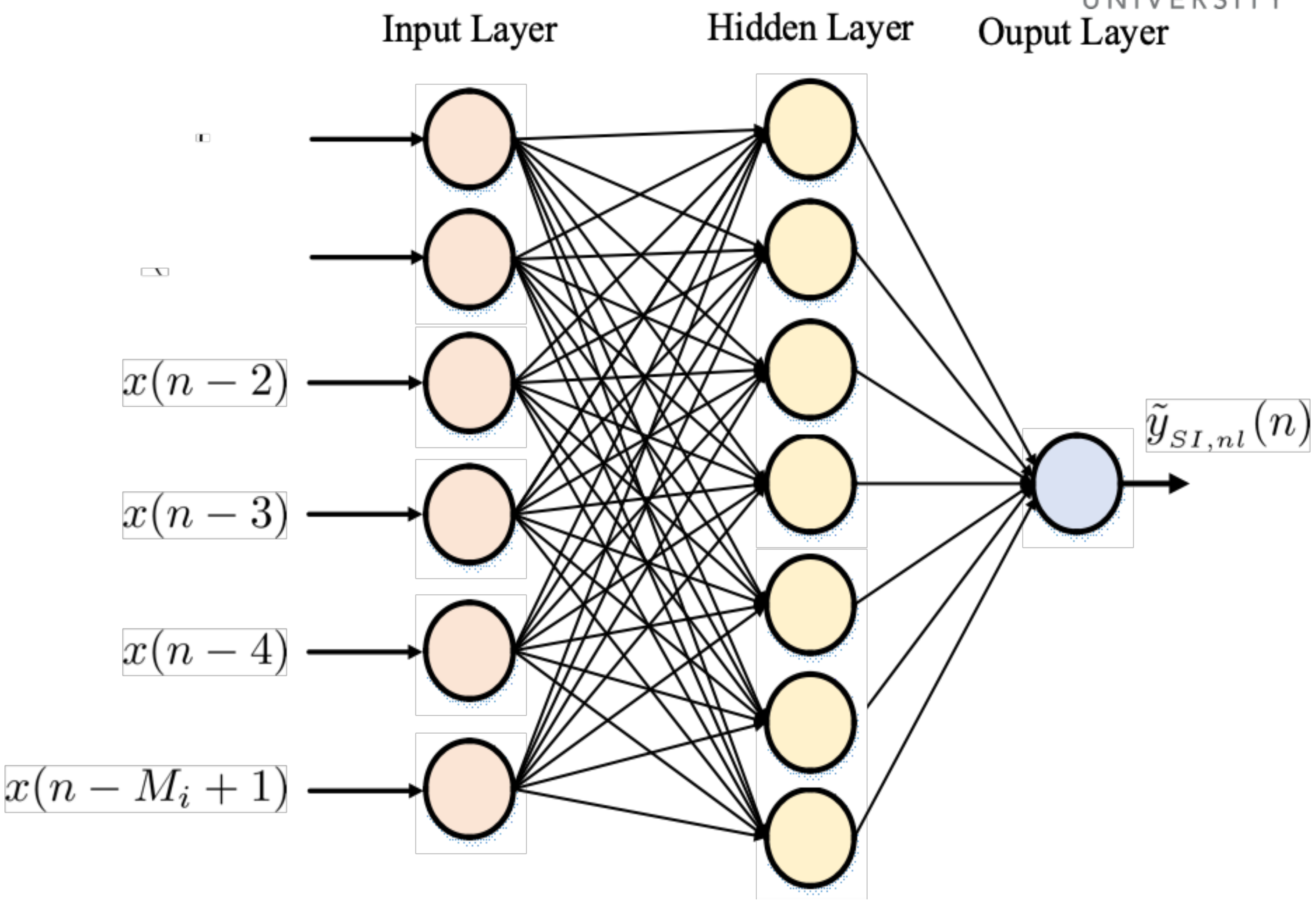
- A. T. Kristensen, A. Burg, and A. Balatsoukas-Stimming, “Advanced machine learning techniques for self-interference cancellation in full-duplex radios,” in *Proc. Asilomar Conf. on Signals, Systems and Computers*, Nov. 2019.

Recent NN-based SIC methods

- **Complex-valued TDNN (CV-TDNN)**

- Suitable candidate for SIC as it employs complex-valued inputs, which is the case in the signal processing in communication systems

- CV-TDNN significantly reduces the number of network parameters without affecting the cancellation performance



CV-TDNN architecture.

- A. T. Kristensen, A. Burg and A. Balatsoukas-Stimming, “Advanced machine learning techniques for self-interference cancellation in full-duplex radios,” in *Proc. Asilomar Conf. on Signals, Systems and Computers*, Nov. 2019.

Achieved Results

➢ Dataset Specifications

	Parameter	Value
Full-duplex testbed	Type of modulation	QPSK-modulated OFDM
	Passband bandwidth	**10 MHz**
	Number of carriers	1024
	Sampling frequency	**20 MHz**
	Average transmit power	**10 dBm**
	Passive analog suppression	53 dB
	Dataset size	**20,480 samples**

- A. T. Kristensen, A. Burg and A. Balatsoukas-Stimming, “Advanced machine learning techniques for self-interference cancellation in full-duplex radios,” in *Proc. Asilomar Conf. on Signals, Systems and Computers*, Nov. 2019.

Achieved Results

MEMORIAL UNIVERSITY

➢ NN parameters

Goal: achieve similar SIC as the polynomial canceler (P=5) with reduced complexity

Simulation parameters of RV-TDNN, RNN, and CV-TDNN.

Parameter	RV-TDNN	RNN	CV-TDNN
Loss Function	MSE	MSE	MSE
Activation Function	ReLU	Tanh	CReLU
Optimizer	Adam	Adam	Adam
Learning Rate	0.005	0.0025	0.0045
Batch Size	22	158	62
Number of Epochs	50	50	50
Validation Split	0.1	0.1	0.1
Number of seeds	20	20	20
M_i	13	-	13

ReLU: rectified linear unit; **CReLU:** complex ReLU; **Adam:** adaptive moment estimation; **MSE:** mean squared error.

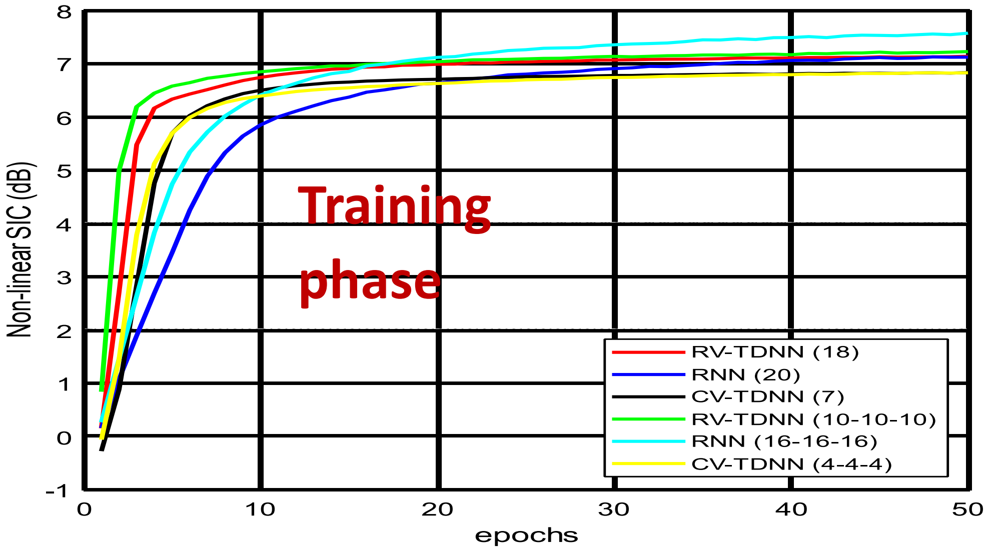
Note: The achieved results above and in the following slides are obtained using the public dataset available at https://github.com/abalatsoukas/fdnn.

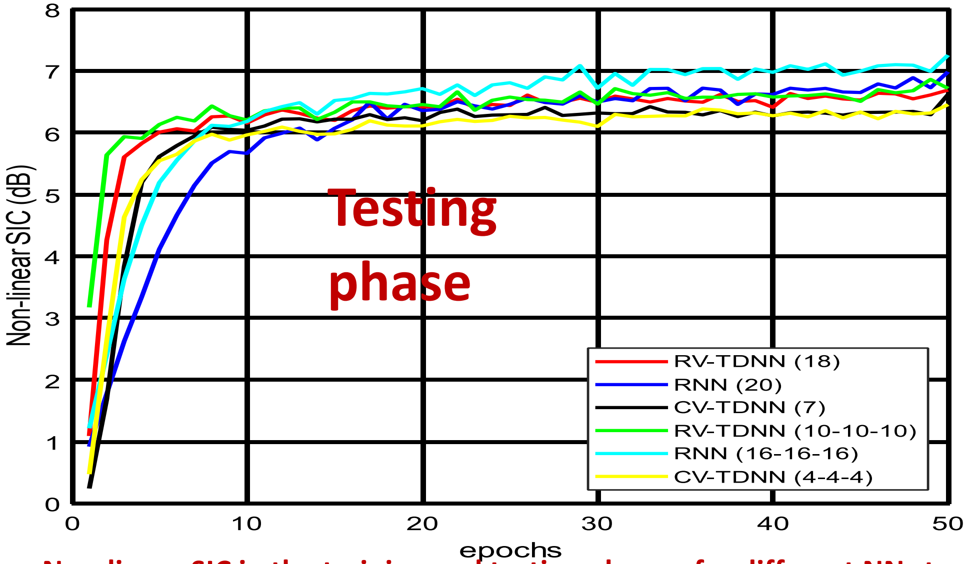
Achieved Results

$$\xi_{dB} = 10\log_{10}\left(\frac{\sum_{n=0}^{K-1}\left|y_{SI}(n)\right|^2}{\sum_{n=0}^{K-1}\left|y_{SI}(n)-\tilde{y}_{SI}(n)\right|^2}\right)$$

MEMORIAL
ERSITY

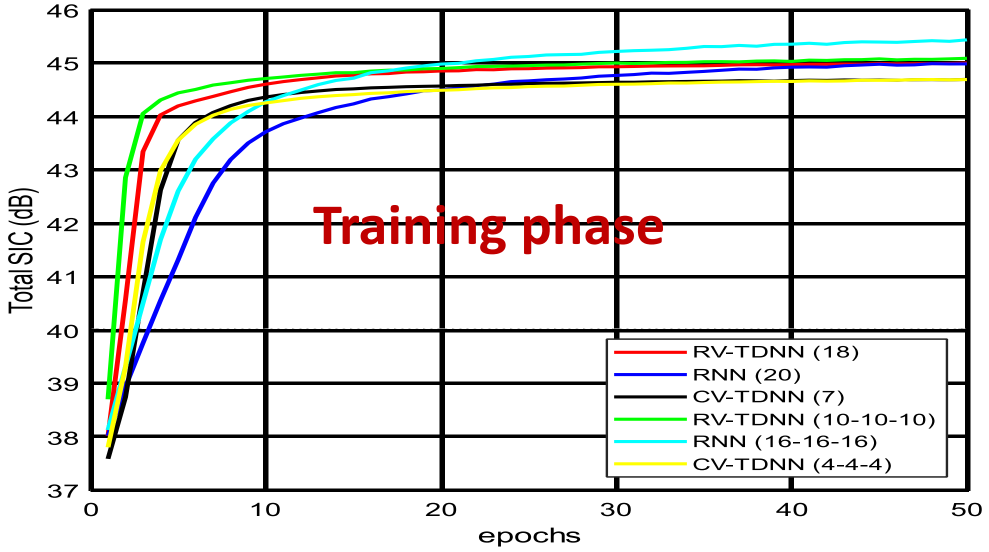
➢Non-linear SIC

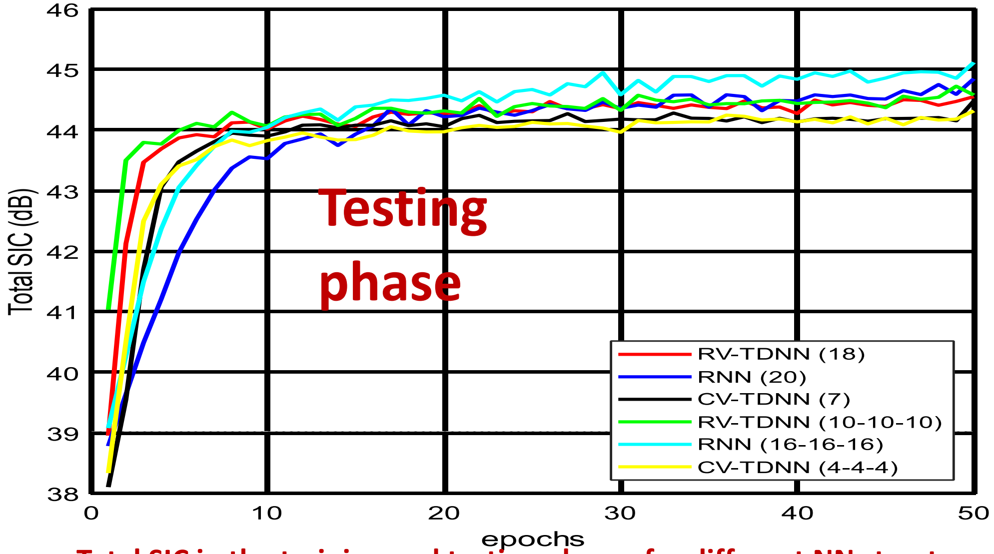




Non-linear SIC in the training and testing phases for different NN structures.

➢Total SIC





Total SIC in the training and testing phases for different NN structures.

Note: These are replicas of the MSE curves.

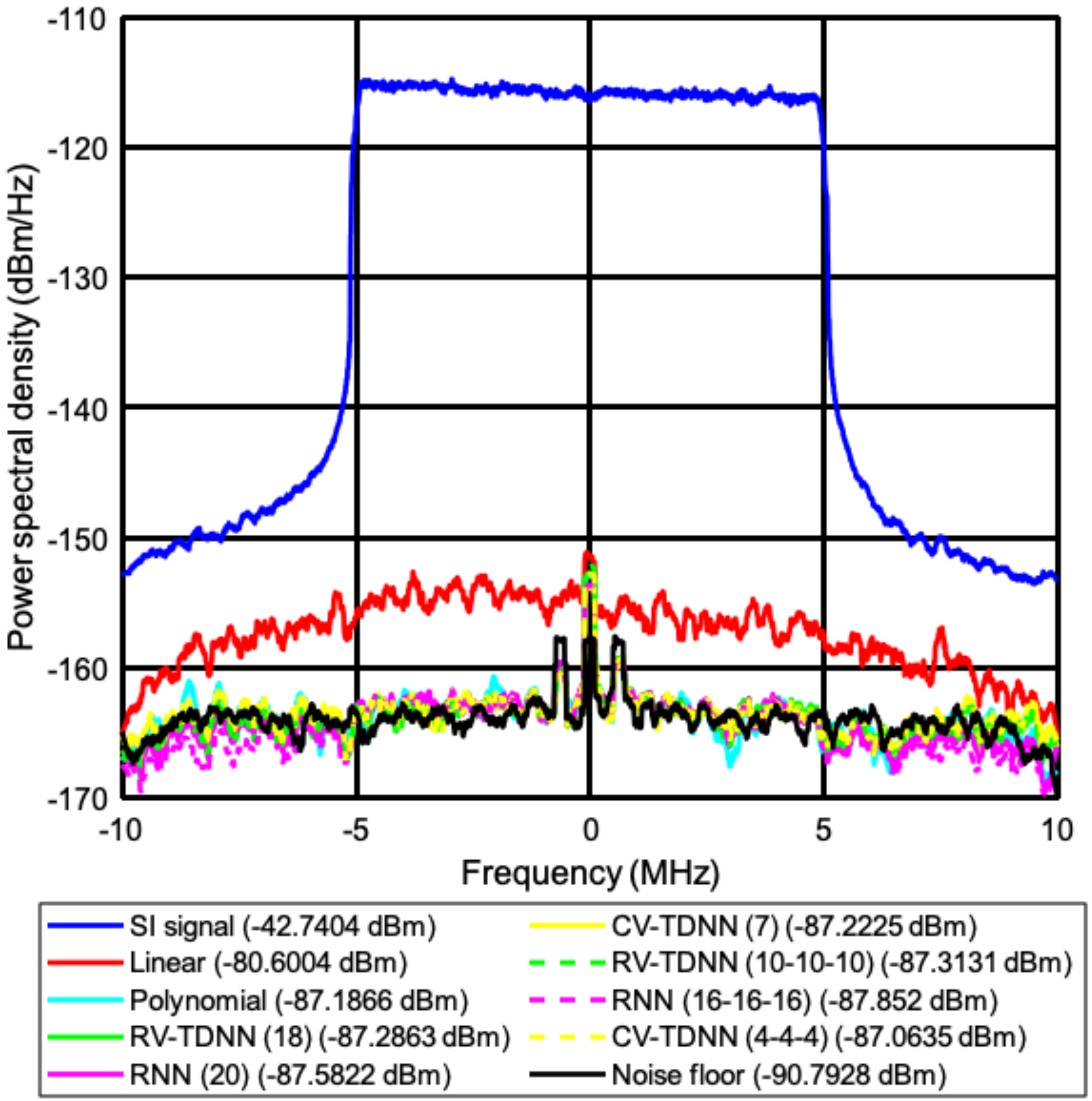
Achieved Results

➢PSD performance

Gap to noise floor.

Network	Canc. (dB)	Gap to Noise Floor (dB)
RV-TDNN (18)	44.76	3.50
RNN (20)	44.94	3.21
CV-TDNN (7)	44.50	3.57
RV-TDNN (10-10-10)	44.73	3.48
RNN (16-16-16)	45.27	2.94
CV-TDNN (4-4-4)	44.63	3.73

- The SI signal is suppressed close to the Rx noise level



PSD curves for NN-based cancelers compared to the polynomial canceler (P = 5).

Achieved Results

MEMORIAL UNIVERSITY

➢ **Complexity reduction compared to the polynomial model at *P* = 5**

Total SIC and complexity of different NN-based cancelers.

Network Structure	SIC (dB)	Total # Parameters	Total # FLOPS	% Parameters	% FLOPs
Polynomial (*P* = 5)	44.45	312	1558	-	-
RV-TDNN (18)	44.76	550	1156	+76.28%	-25.80%
RNN (20)	44.94	528	1210	+69.23%	-22.34%
CV-TDNN (7)	44.50	238	1166	**-23.72%**	**-25.16%**
RV-TDNN (10-10-10)	44.73	538	1120	+72.44%	-28.11%
RNN (16-16-16)	45.27	1420	3106	+355.13%	+99.36%
CV-TDNN (4-4-4)	44.63	228	1106	**-26.92%**	**-29.01%**

Note: Number of FLOPs = number of RV multiplications and additions for linear and non-linear cancellations

- A CV multiplication: 3 RV multiplications & 5 RV additions
- A CV addition: 2 RV additions

Conclusion

- **RNN**:
  - higher number of parameters than the polynomial based canceler
  - more training epochs to converge
- **RV-TDNN**:
  - higher number of parameters than the polynomial based canceler
  - less training epochs to converge
- **CV-TDNN**:
  - significantly reduces the number of FLOPs and parameters than the polynomial model
- **We conclude that**:
  - **RNN** structures are not practical candidates for SIC
  - **CV-TDNN** can be a suitable candidate for SIC from the FLOPs and parameters reduction perspective

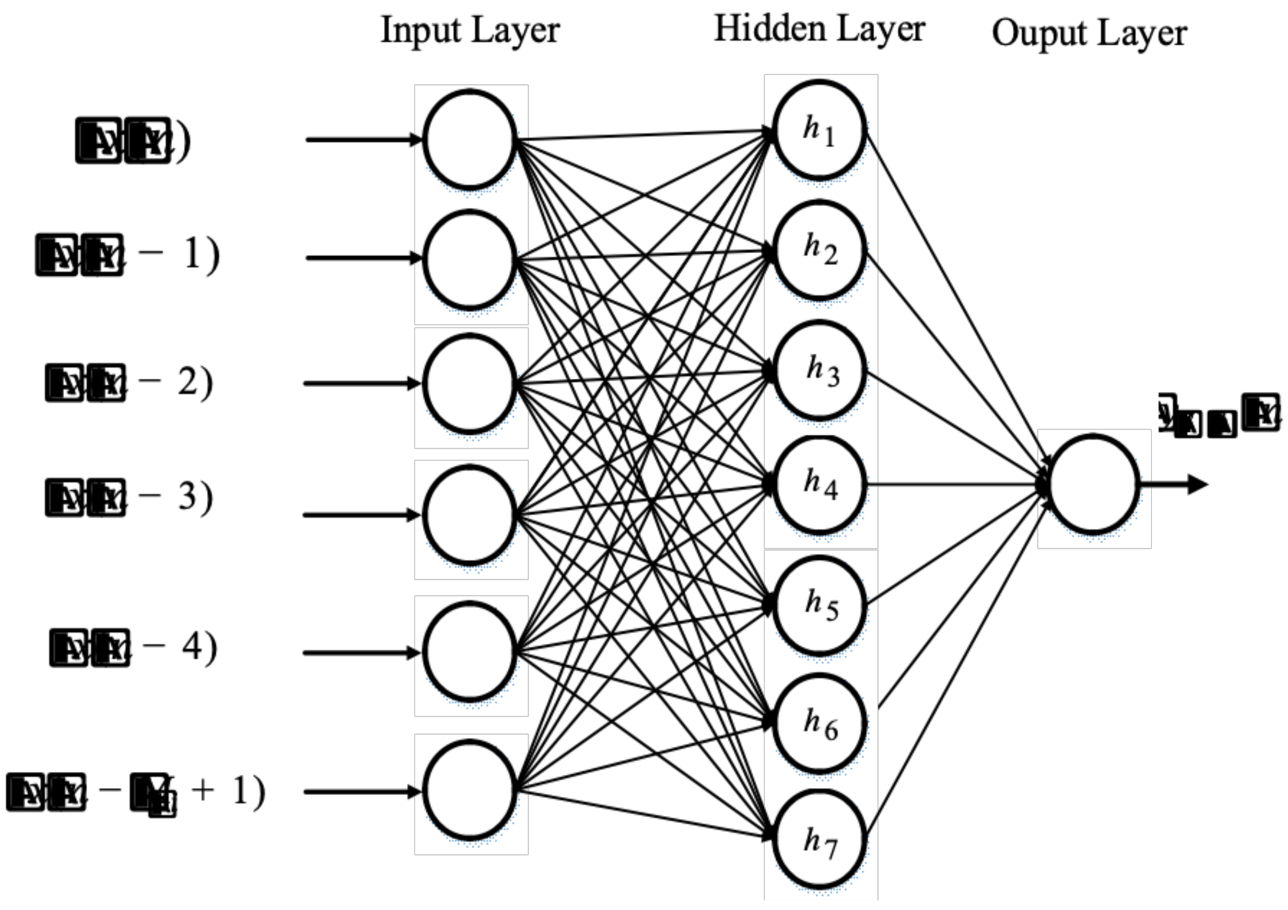
Idea 1: Grid-based NN Structures

- Design 1: Ladder-wise grid structure (LWGS)
- Design 2: Moving-window grid structure (MWGS)

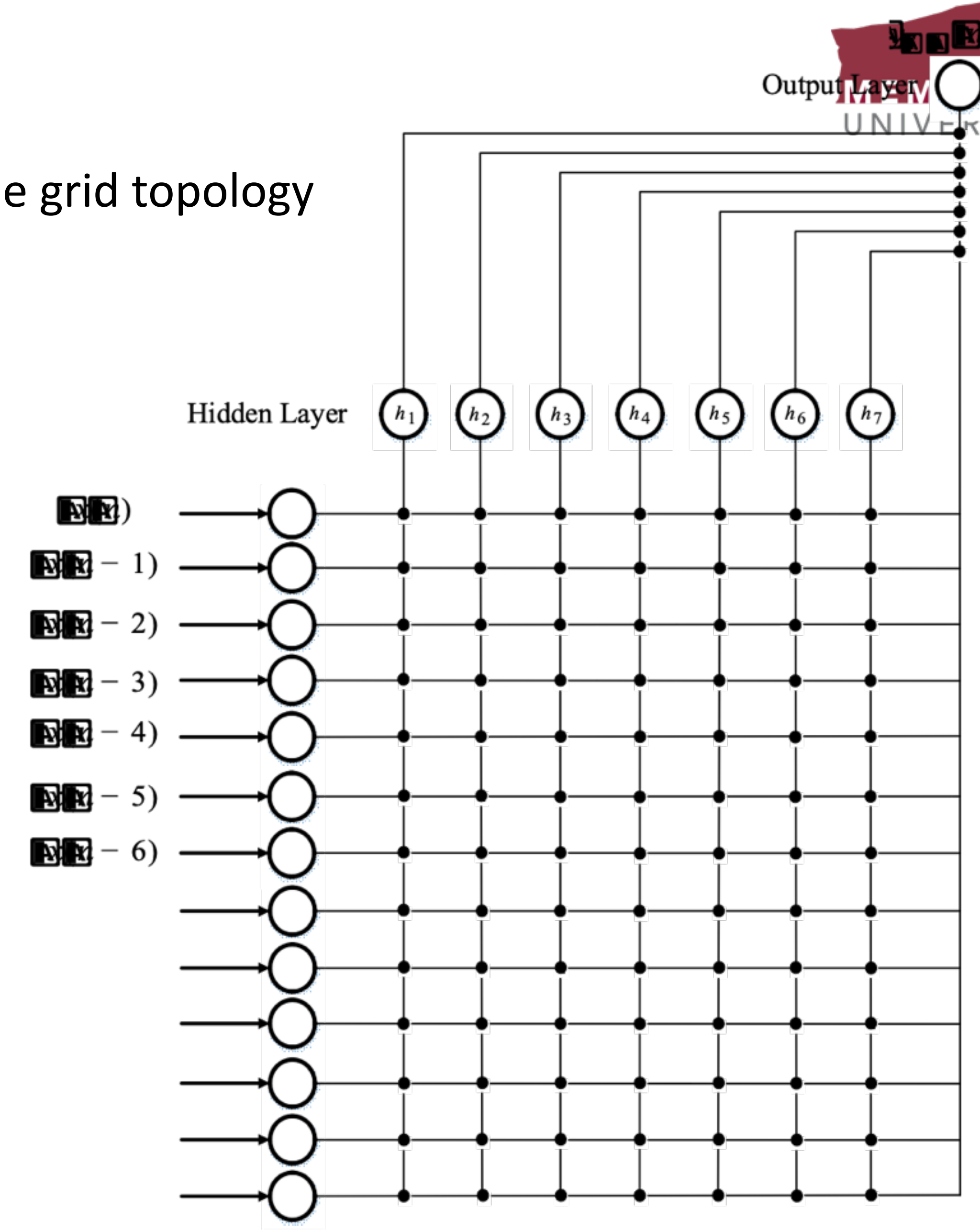
- M. Elsayed, A. A. A. El-Banna, O. A. Dobre, W. Shiu, and P. Wang, “Low complexity neural network structures for self-interference cancellation in full-duplex radio,” *IEEE Commun. Lett.*, vol. 25, no. 1, pp. 181-185, Jan. 2021.

Overview of Grid-based Structures

- The fully-connected CV-TDNN can be represented using the grid topology

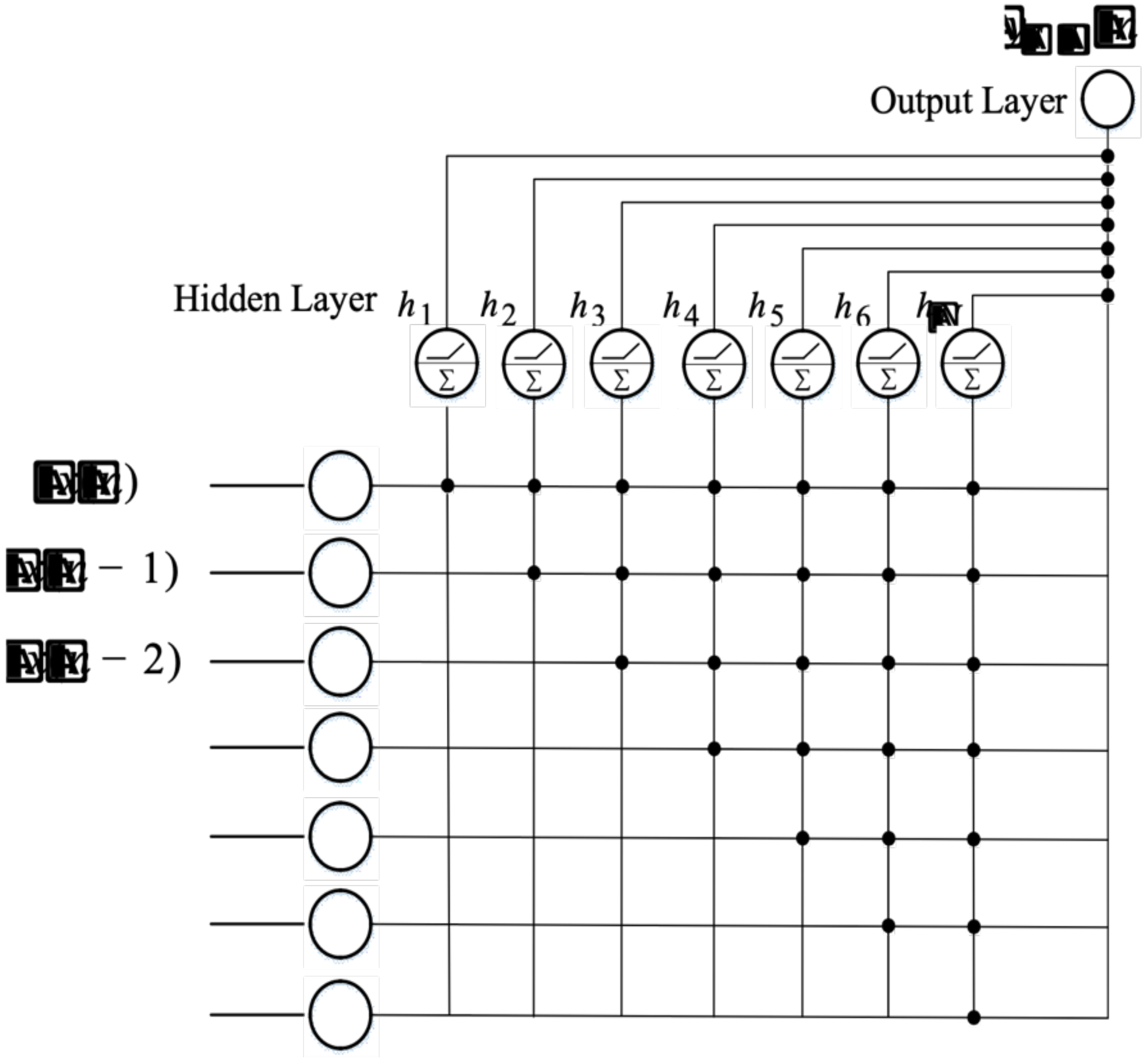


Fully-connected CV-TDNN.

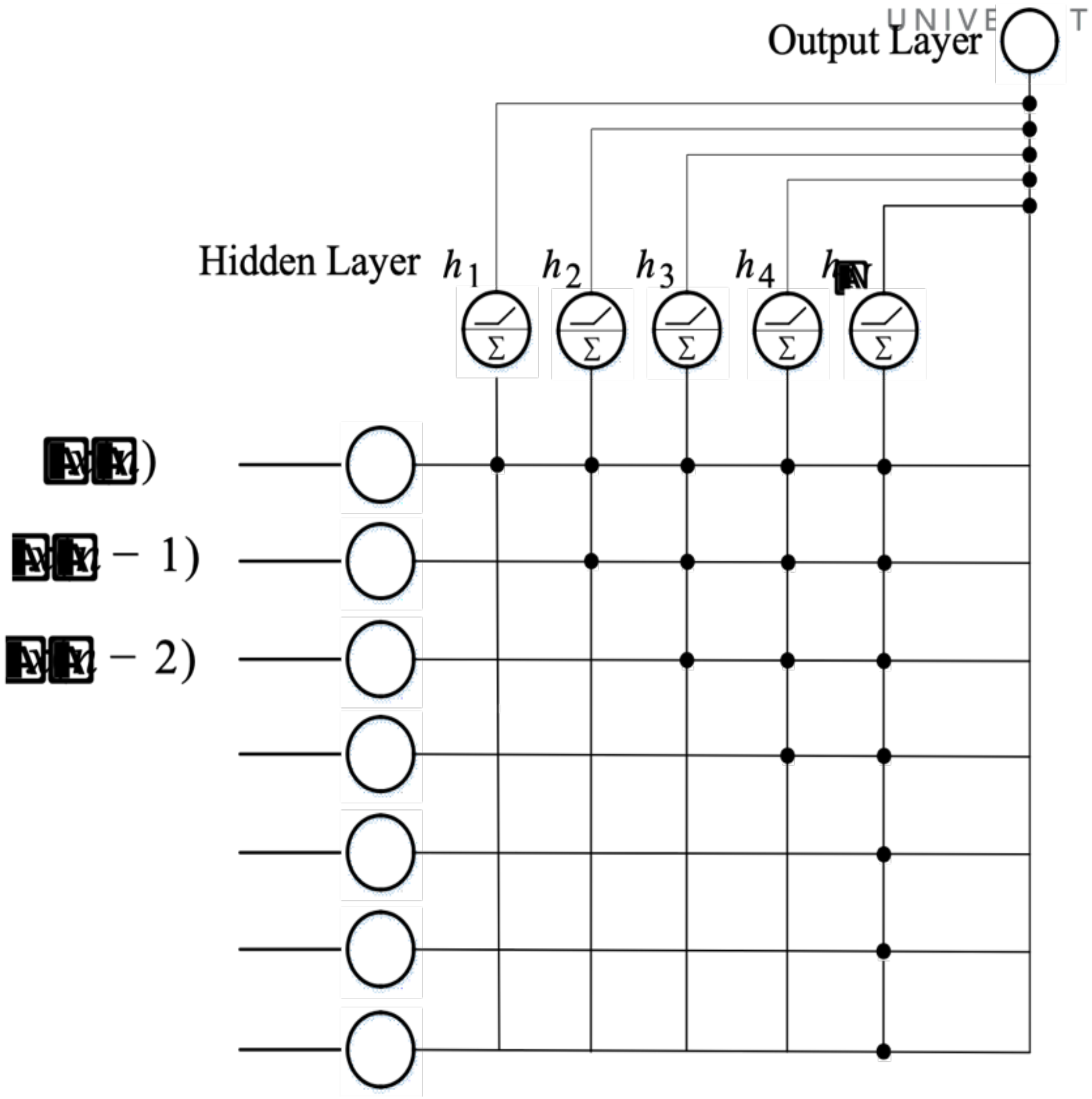


Grid representation of the fully-connected CV-TDNN.

Design 1: Ladder-Wise Grid Structure (LWGS)



LWGS (complete ladder).

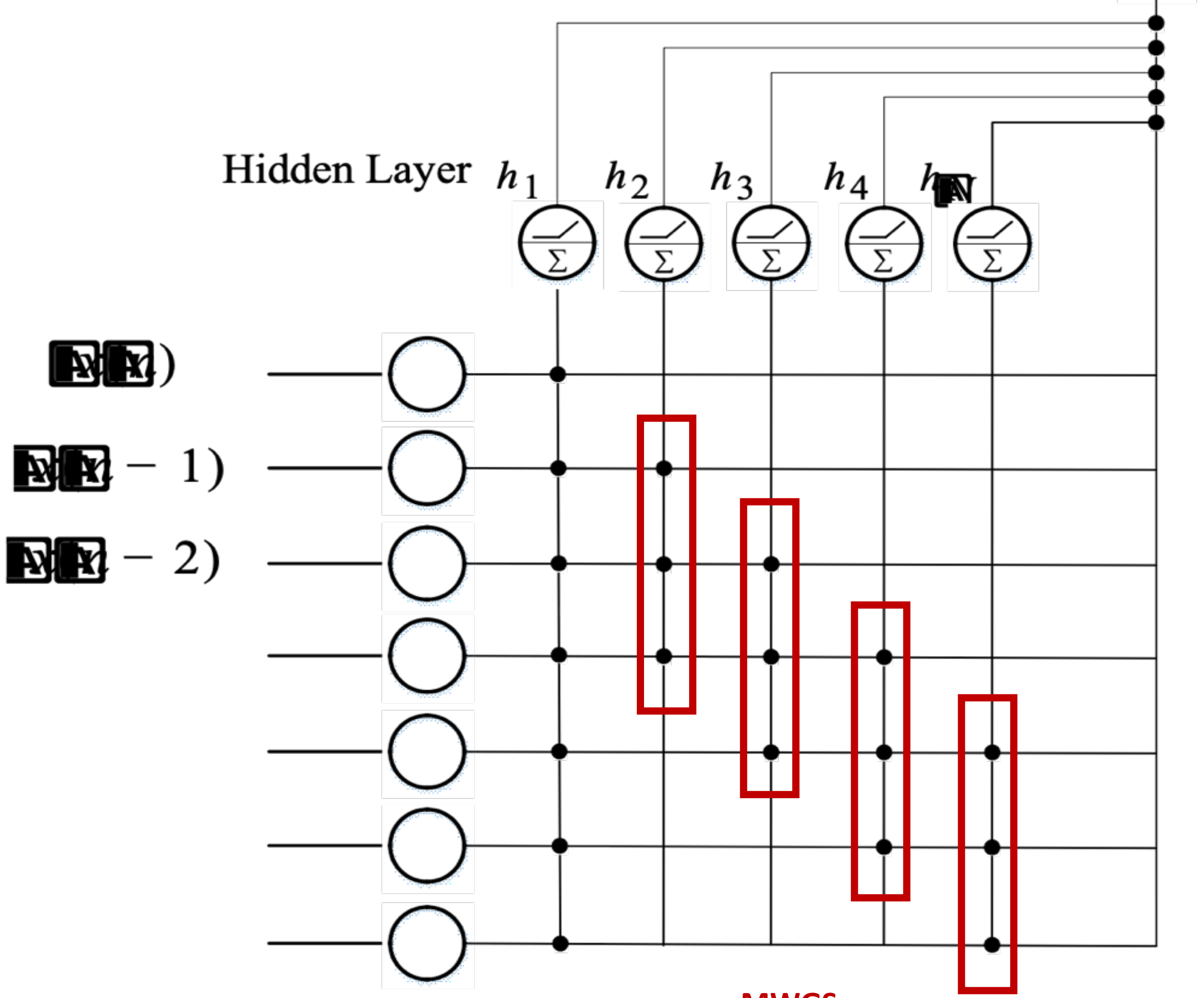


LWGS (shorter length of ladder).

Design 2: Moving-Window Grid Structure (MWGS)

Moving window procedure: recognized as an effective method for time series prediction

- **MWGS**
  - The input samples learned by neurons are varied based on a fixed-length moving window technique
  - All input samples are passed to the first neuron
  - The other neurons assist in learning the memory effect by considering the windowed data
  - Sliding the window over different samples allows to consider all buffered samples of the input signal



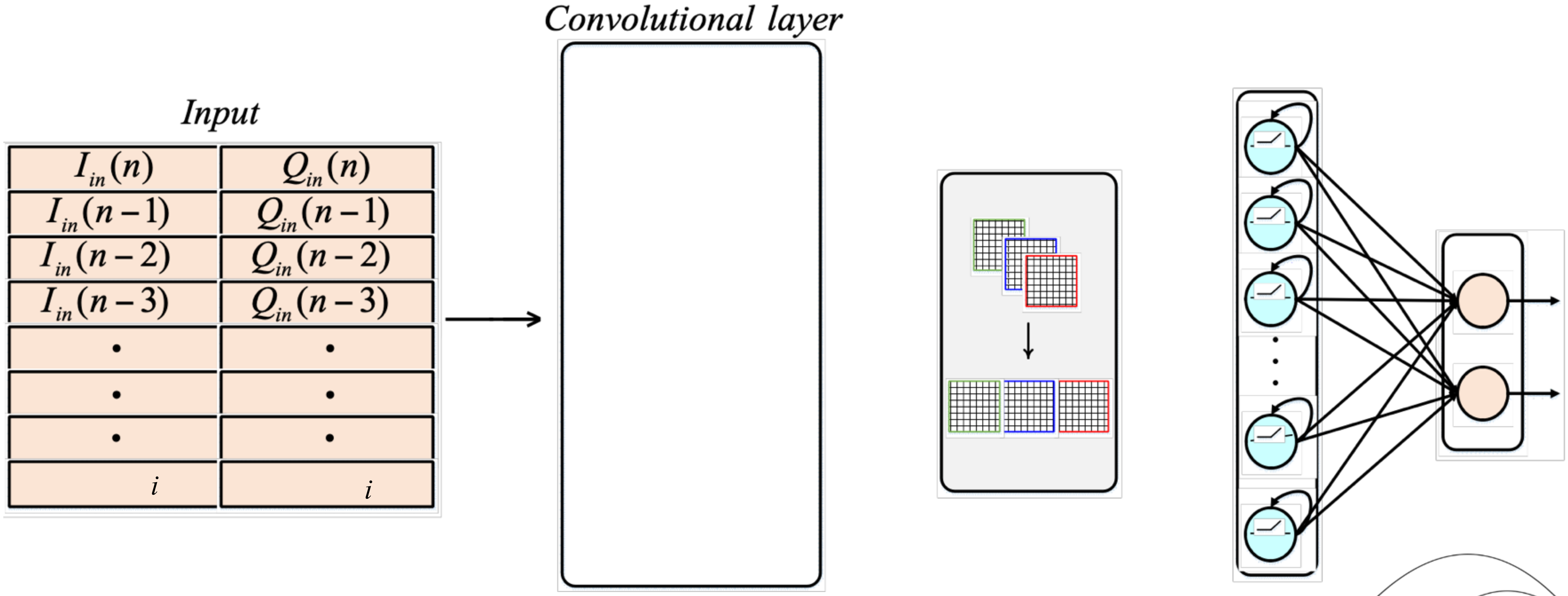
MWGS.

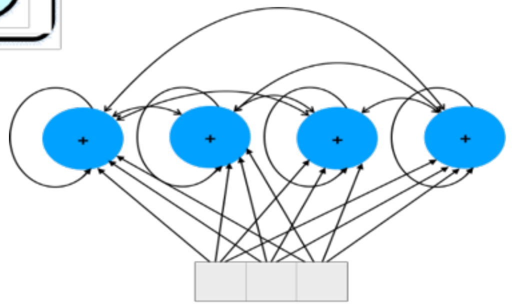
Idea 2: Hybrid-Layers NN Structures

- Design 1: Hybrid convolutional-recurrent NN (HCRNN)
- Design 2: Hybrid convolutional-recurrent-dense NN (HCRDNN)

- M. Elsayed, A. A. El-Banna, O. A. Dobre, W. Shiu, and P. Wang, "Hybrid-layers neural network architectures for modeling the self-interference in full-duplex systems," *IEEE Transactions on Vehicular Technology*, vol. 71, issue 6, pp. 6291-6307, June 2022.

Design 1: Hybrid Convolutional Recurrent NN (HCRNN)

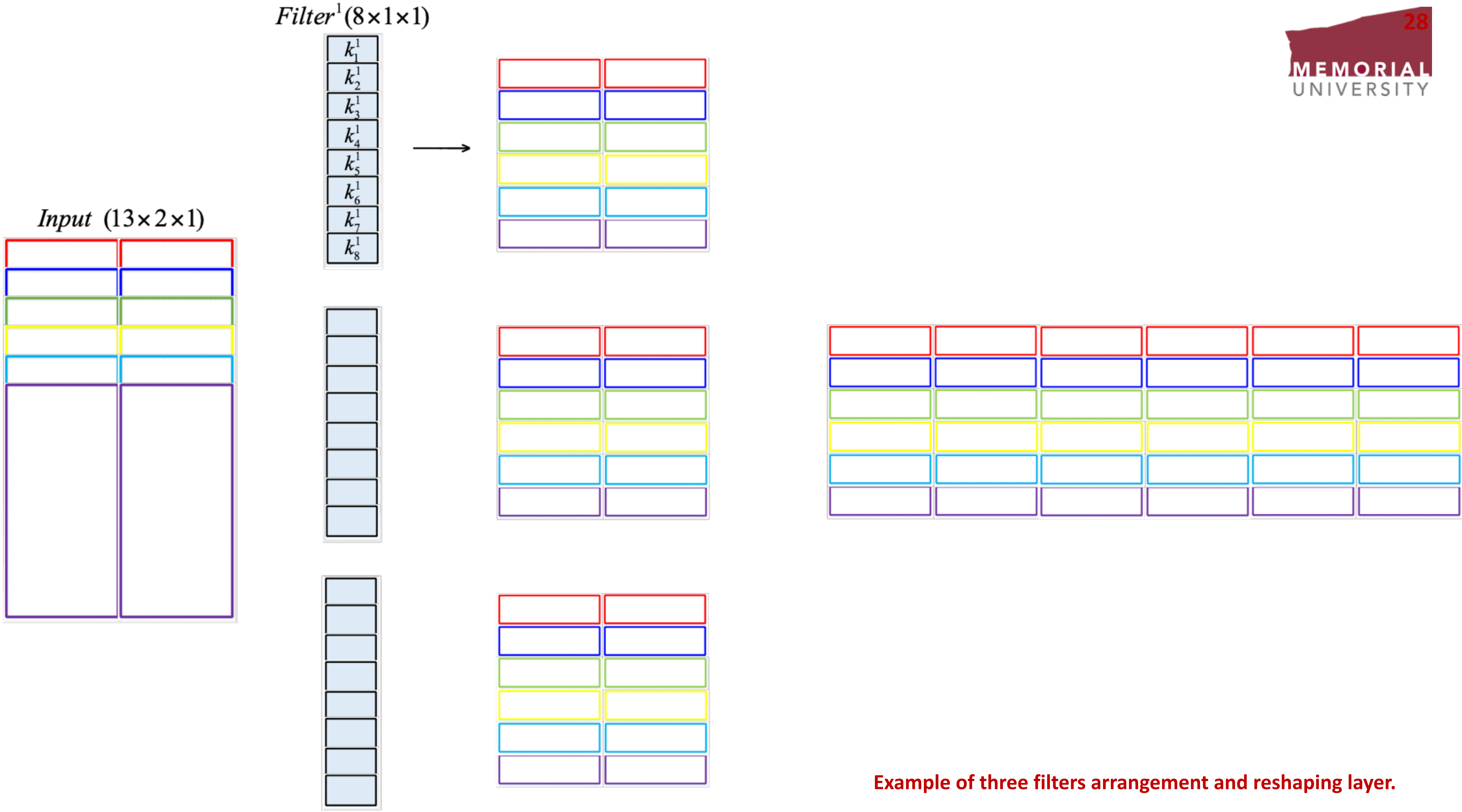




HCRNN structure.

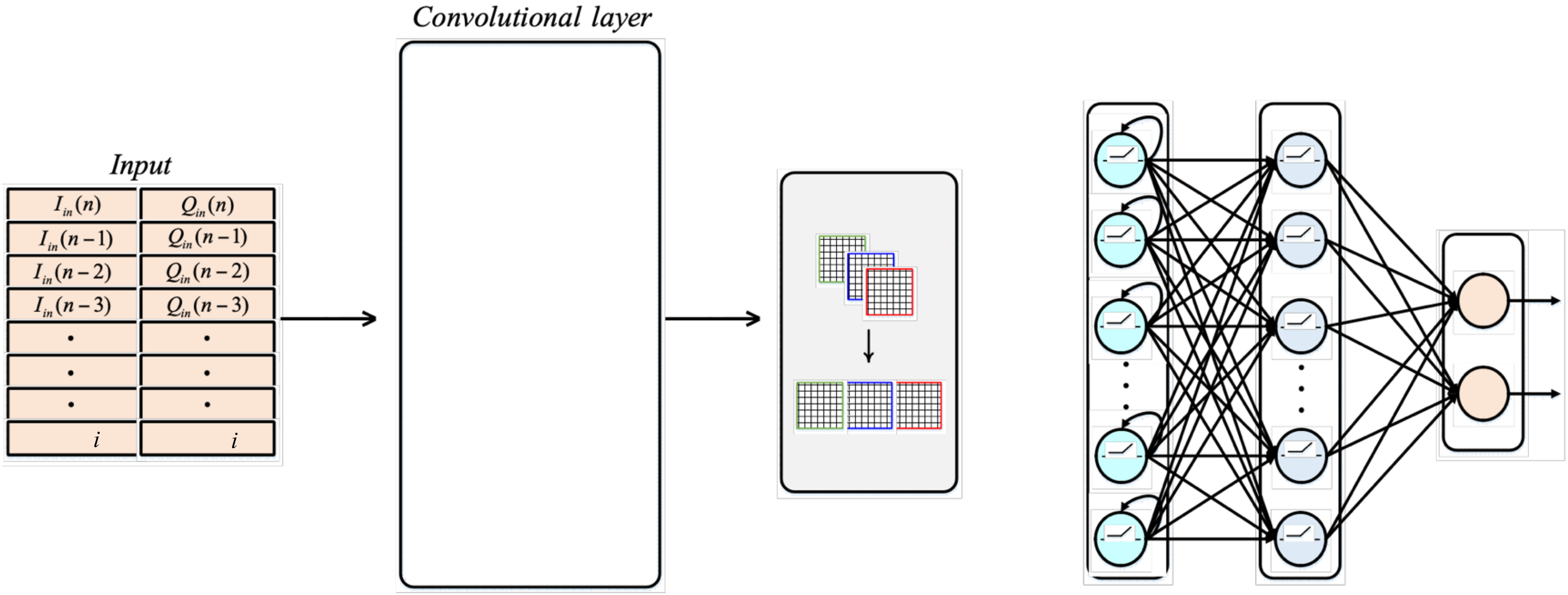
Notes: - Recurrent layer: feedback connections from each neuron to all neurons

- $I'_{out}(n)$ and $Q'_{out}(n)$ are the estimated I/Q components of the SI signal $y_{SI,nl}(n)$



Example of three filters arrangement and reshaping layer.

Design 2: Hybrid convolutional recurrent dense NN (HCRDNN)



HCRDNN structure.

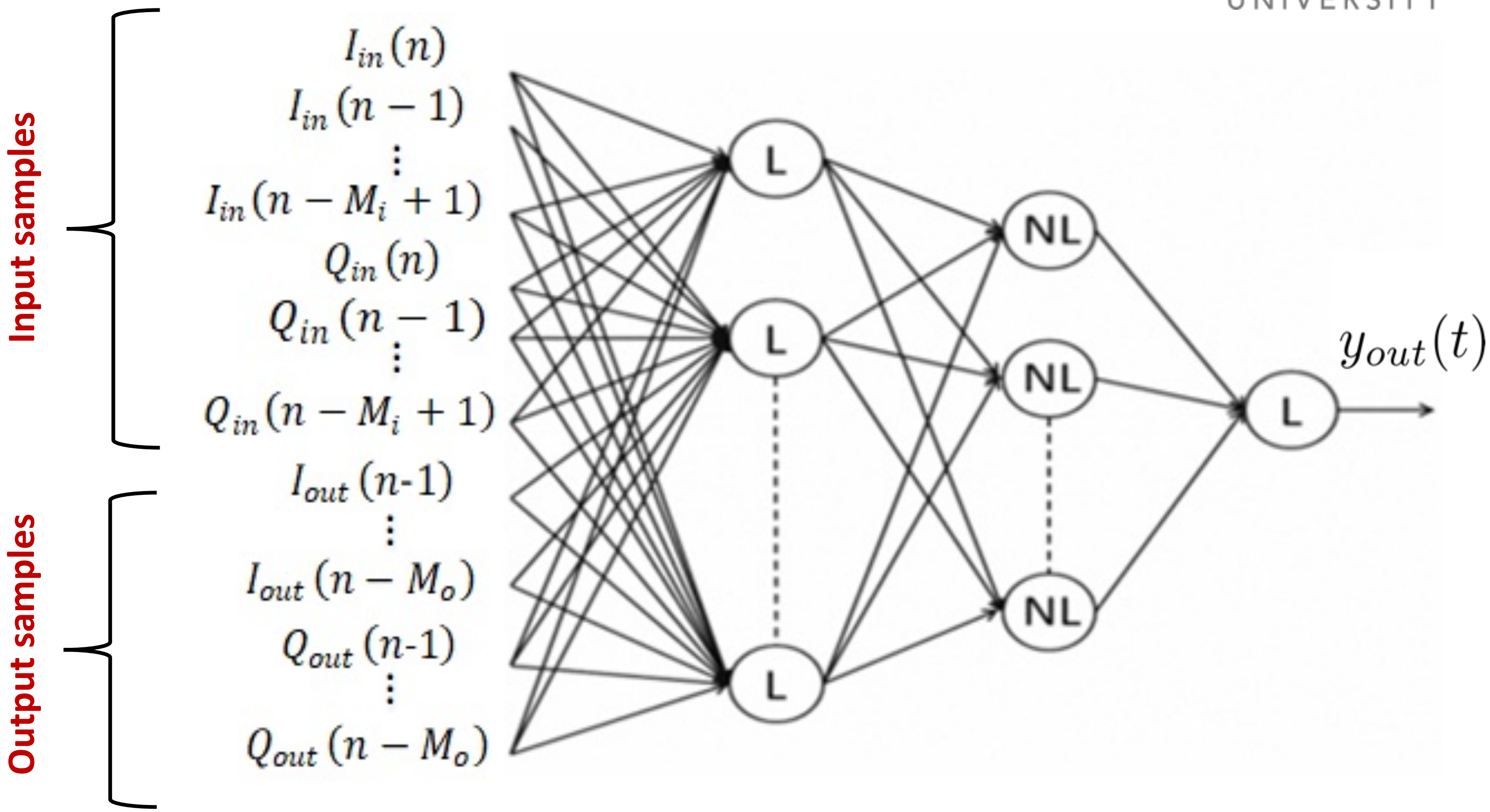
Idea 3: Dual-Neurons NN Structures

- RV-2HLNN structure
- Idea of the proposed DN-ℓHLNN

- M. Elsayed, A. A. A. El-Banna, O. A. Dobre, W. Shiu, and P. Wang, "Full-duplex self-interference cancellation using dual-neurons neural networks," *IEEE Commun. Lett.*, vol. 26, no. 3, pp. 557-561, March 2022.

RV-2HLNN Structure

- An RV-FFNN known as a two-hidden layer NN (2HLNN) is introduced to model the non-linearity of memory-based systems such as the Doherty RF PA. **This is an RV NN**
- The RV-2HLNN is constructed based on behavioral modeling of the PA where delayed versions of the input and output samples are utilized as attributes to the input layer
- It is worth mentioning that for the RF PA, the RV-2HLNN significantly outperforms the RV-TDNN



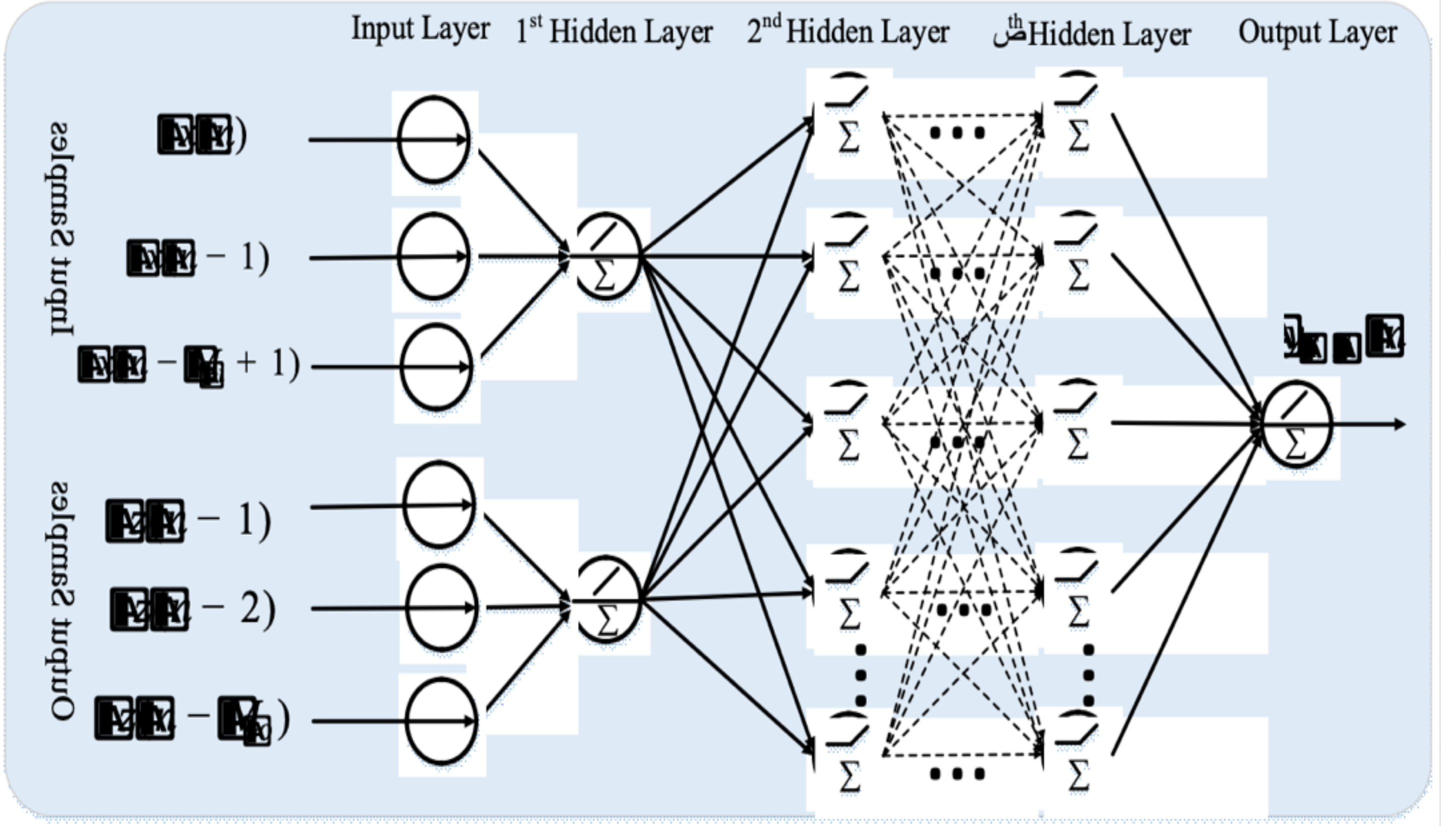
RV-2HLNN-based model for PA.

- M_i and M_o designate the memory depth attributed to the input and output signals, respectively

- F. Mkadem and S. Boumaiza, "Physically inspired neural network model for RF power amplifier behavioral modeling and digital predistortion," *IEEE Trans. Microw. Theory Technol.*, Apr. 2011.

Idea of the Proposed DN-ℓHLNN

- The dual neurons-ℓ hidden layers neural network (DN-ℓHLNN) employs the CV framework
- **The first hidden layer:**
  - The input units are **not fully connected**
  - Uses two neurons to recognize the memory effect of the input and output signals separately, while reducing the required number of network's parameters (e.g., weights and biases)
  - The activation functions are linear functions
- **The other hidden layers (i.e., 2nd,3rd, …, ℓth):**
  - Approximate the non-linearity induced by the various components of the FD transceiver
  - The activation functions are non-linear functions



Proposed DN-ℓHLNN structure.

$$z(n) = y_{SI}(n) - \tilde{y}_{SI,lin}(n)$$

Achieved Results

Achieved Results

- Optimum settings for training the NN architectures based on hyperparameter tuning

NN model parameters.

Parameter	Value	
	RV structures	CV structures
Loss function	MSE	MSE
Learning rate	0.005	0.0045
Batch size	62	62
Activation function	ReLU	Complex-ReLU
Optimizer	Adam	Adam
Number of epochs	50	50
Validation split	0.1	0.1
Number of seeds	15	15

Notes: - HCRDNN 1 and HCRDNN 2 are trained using a 158 batch size.
- The achieved results are obtained using the **public dataset** available at https://github.com/abalatsoukas/fdnn.

MEMORIAL UNIVERSITY

Achieved Results

Other parameters for the NN models.

Structure	CV-TDNN	LWGS	MWGS	HCRNN	HCRDNN 1	HCRDNN 2
# Neurons in the hidden layer	7	9	12			
Window size			5			
# Filters				3	2	3
Filter size				12×1×1	12×1×1	12×1×1
# Rec. neurons				9	7	5
# Dense neurons					11	12

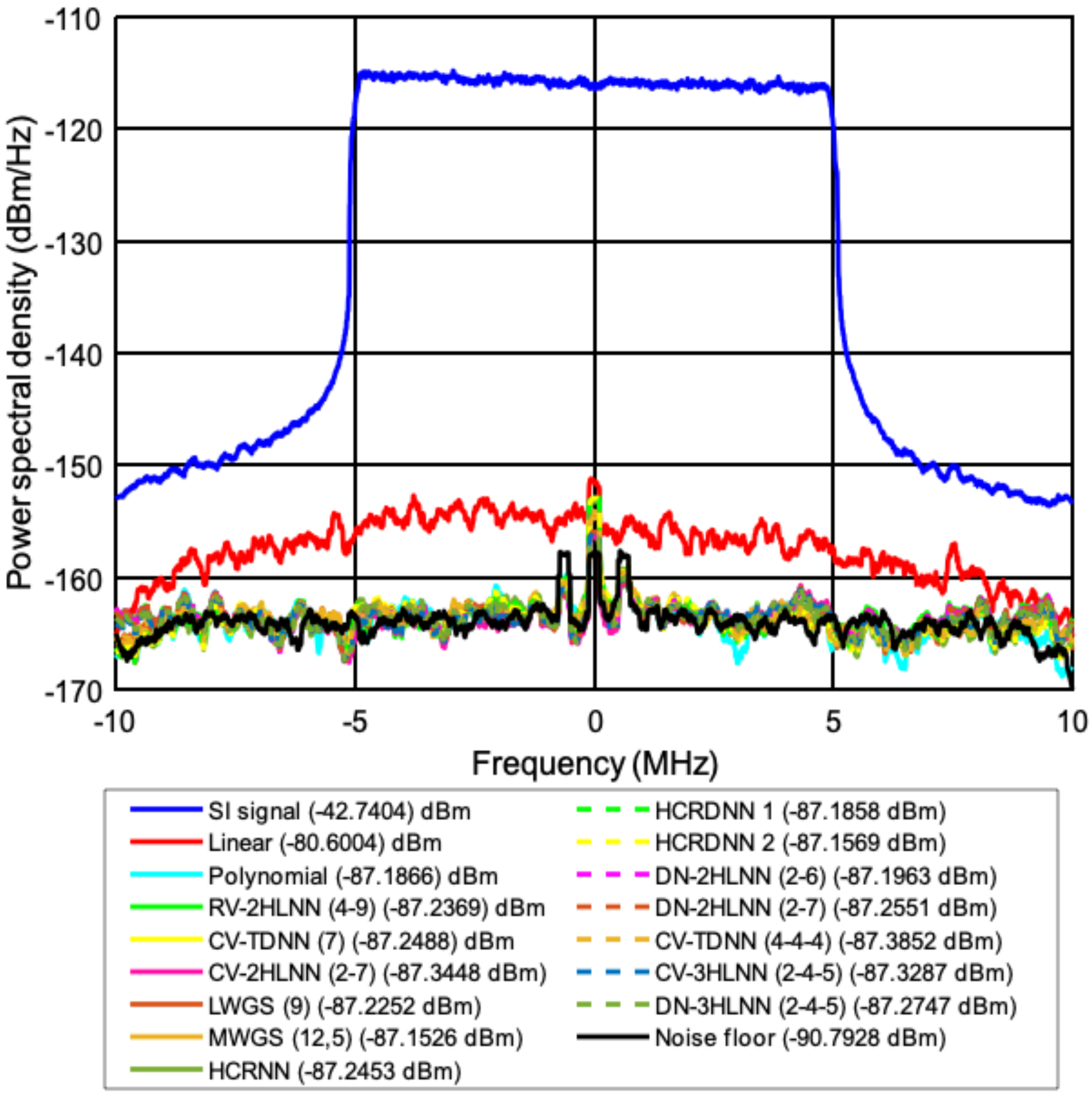
Note: These structures achieve a similar SIC performance with reduced computational complexity (from each of the proposed NN structures)

Achieved Results

➢ **PSD performance**

Gap to noise floor for the of the best NN candidates.

Network	Canc. (dB)	Gap to noise floor (dB)
RV-2HLNN (4-9)	44.50	3.56
CV-TDNN (7)	44.50	3.54
CV-2HLNN (2-7)	44.58	3.45
LWGS (9)	44.48	3.57
MWGS (12,5)	44.40	3.64
HCRNN	44.50	3.55
HCRDNN 1	44.44	3.61
HCRDNN 2	44.41	3.64
DN-2HLNN (2-6)	44.44	3.60
DN-2HLNN (2-7)	44.50	3.54
CV-TDNN (4-4-4)	44.63	3.41
CV-3HLNN (2-4-5)	44.57	3.46
DN-3HLNN (2-4-5)	44.51	3.52



PSD curves for the best NN candidates compared to the polynomial canceler (P = 5).

SUMMARY of RESULTS (Dataset #1)

- Results with the public dataset (https://github.com/abalatsoukas/fdnn)

Total SIC of different NN-based cancelers (dataset #1).

Canceler type	Network	Total average SIC (dB)	Linear SIC (dB)	Non-linear SIC (dB)	Gap to Rx Noise Floor (dB)	Linear Canceler Complexity		NN Model Complexity		Total Complexity		Complexity Reduction to Polynomial (P=5)	
						# Par.	# FLOPS	# Par.	# FLOPS	# Par.	# FLOPS	# Par.	# FLOPS
Baseline	Polynomial (P=5)	44.45		6.59	3.61			-	-	312	1558	-	-
Real-valued NN	RV-2HLNN (4-9)	44.50		6.64	3.56			269	517	295	647	-5.45%	-58.47 %
	HCRNN	44.50		6.64	3.54			203	615	229	745	-26.60%	-52.18 %
	HCRDNN 1	44.58		6.72	3.45			222	570	248	700	-20.51%	-55.07 %
	HCRDNN 2	44.48		6.62	3.57			197	595	223	725	-28.53%	-53.47 %
	CV-TDNN (7)	44.40		6.54	3.64			212	1036	238	1166	-23.72%	-25.16 %
	CV-2HLNN (2-7)	44.50		6.64	3.55			162	766	188	896	**-39.74%**	**-42.49 %**
	LWGS (9)	44.44	37.86	6.58	3.61	26	128	136	652	162	782	**-48.08%**	**-49.81 %**
	MWGS (12,5)	44.41		6.55	3.64			186	896	212	1026	-32.05%	-34.15 %

Note: Results are compared to polynomial canceler with ***P* = 5**, which requires 1558 FLOPs and 312 parameters to achieve same cancellation.

SUMMARY of RESULTS (Dataset #2)

MEMORIAL UNIVERSITY

- Results with a second public dataset (https://github.com/abalatsoukas/CSI-full-duplex.
  Average transmit power is **32 dBm** (instead of 10 dBm in the first dataset) and **sampling rate** is **$4F_N$ (instead of $2F_N$)**.

Total SIC of different NN-based cancelers (dataset #2).

Canceler type	Network	Total average SIC (dB)	Linear SIC (dB)	Non-linear SIC (dB)	Gap to Rx Noise Floor (dB)	Linear Canceler Complexity		NN Model Complexity		Total Complexity		Complexity Reduction to Polynomial (P=5)	
						# Par.	# FLOPS	# Par.	# FLOPS	# Par.	# FLOPS	# Par.	# FLOPS
Baseline	Polynomial (P=5)	30.40		11.29	**11.67**			-	-	312	1558	-	-
Real-valued NN	RV-2HLNN (4-9)	33.73		14.62	8.34			269	517	295	647	-5.45%	-58.47 %
	HCRNN	33.87		14.77	8.20			203	615	229	745	-26.60%	-52.18 %
	HCRDNN 1	33.94		14.84	8.13			222	570	248	700	-20.51%	-55.07 %
	HCRDNN 2	34.07		14.96	8.00			197	595	223	725	-28.53%	-53.47 %
	CV-TDNN (7)	34.09		14.98	7.98			212	1036	238	1166	-23.72%	-25.16 %
	CV-2HLNN (2-7)	35.56		16.46	6.51			162	766	188	896	-39.74%	-42.49 %
	LWGS (9)	30.89	19.11	11.79	11.18	26	128	136	652	162	782	-48.08%	-49.81 %
	MWGS (12,5)	34.47		15.36	7.61			186	896	212	1026	32.05%	-34.15 %

In the following slides, we will drop the CV from the CV-2HLNN, CV-3HLNN, etc. structures for the ease of notation.

- F. Jochems and A. Balatsoukas-Stimming, "Non-Linear Self-Interference Cancellation via Tensor Completion," in *Proc. Asilomar Conference on Signals, Systems and Computers*, 2020.

Summary (full data set used for training)

- Polynomial model for SIC
  - Can be accurate for representing the SI
  - Requires high computational complexity

- **NN-based SIC: Model-Centric Approach**: appealing tool to model the SI with lower computational complexity
  - RV-TDNN: lower complexity than the polynomial model
  - RNN: not a practical candidate for SIC
  - CV-TDNN: reduces the number of FLOPs & parameters compared to the polynomial model
  - Our proposed solutions further reduce the complexity

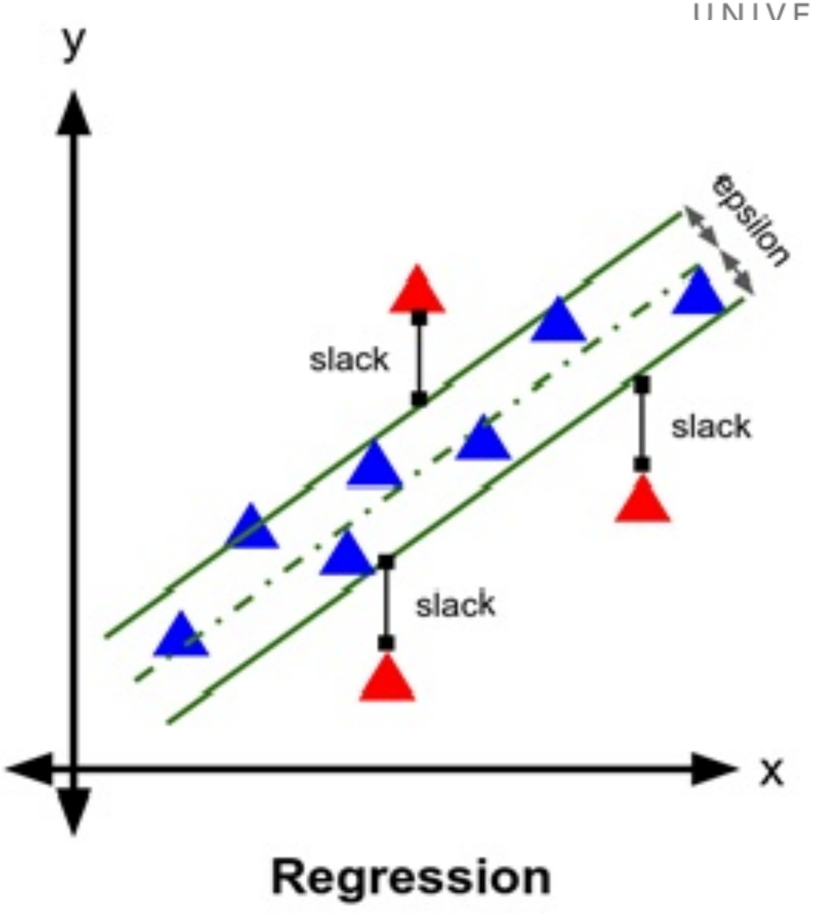
- Using the public datasets
  - Superiority of proposed NN structures vs. polynomial and the existing NN-based cancelers in terms of complexity
  - **DN-2HLNN**: lowest complexity and provides about 60% reduction in the number of network parameters and FLOPs over the polynomial-based canceler
  - Some structures trained for dataset #1 perform reasonably well when applied to dataset #2 (change in the scenario)

Training: with 20,480 samples. What about if the channel changes and we need to train again?

Support Vector Regressor (SVR)-based SIC

Motivation

- **NNs can succumb to various problems, such as:**
  - expensive training cost
  - poor generalization, especially when few examples are available for training
- **Support vector machines are generally very fast to train:**
  - use only a subset of a dataset as training data
  - particularly well suited for problems of complex but small- or medium-sized datasets
- **Main objective:**
  - employ the SVR to model the non-linear SI components in operating scenarios where few data samples are available for training



SVM for regression (SVR).

- A. Géron, Hands-on Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. Sebastopol, CA, USA: O'Reilly, 2017.

Motivation

SVR

Objective: consider the points that are within the decision boundary lines

The best-fit line is the hyperplane that has a maximum number of points

Linear SVR:

Employs a linear kernel

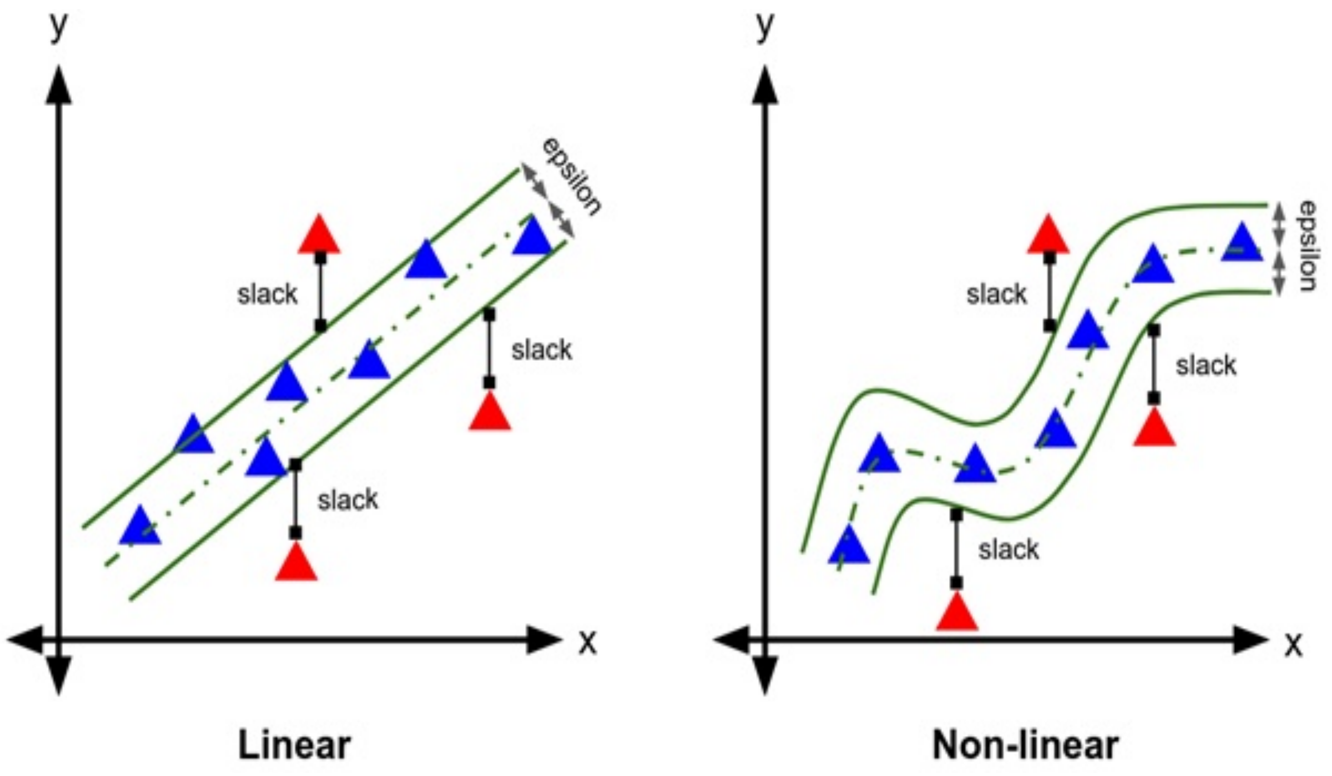
Non-linear SVR:

Employs a non-linear kernel (e.g., RBF)

SVR optimization problem:

Solved using the method of Lagrange multipliers

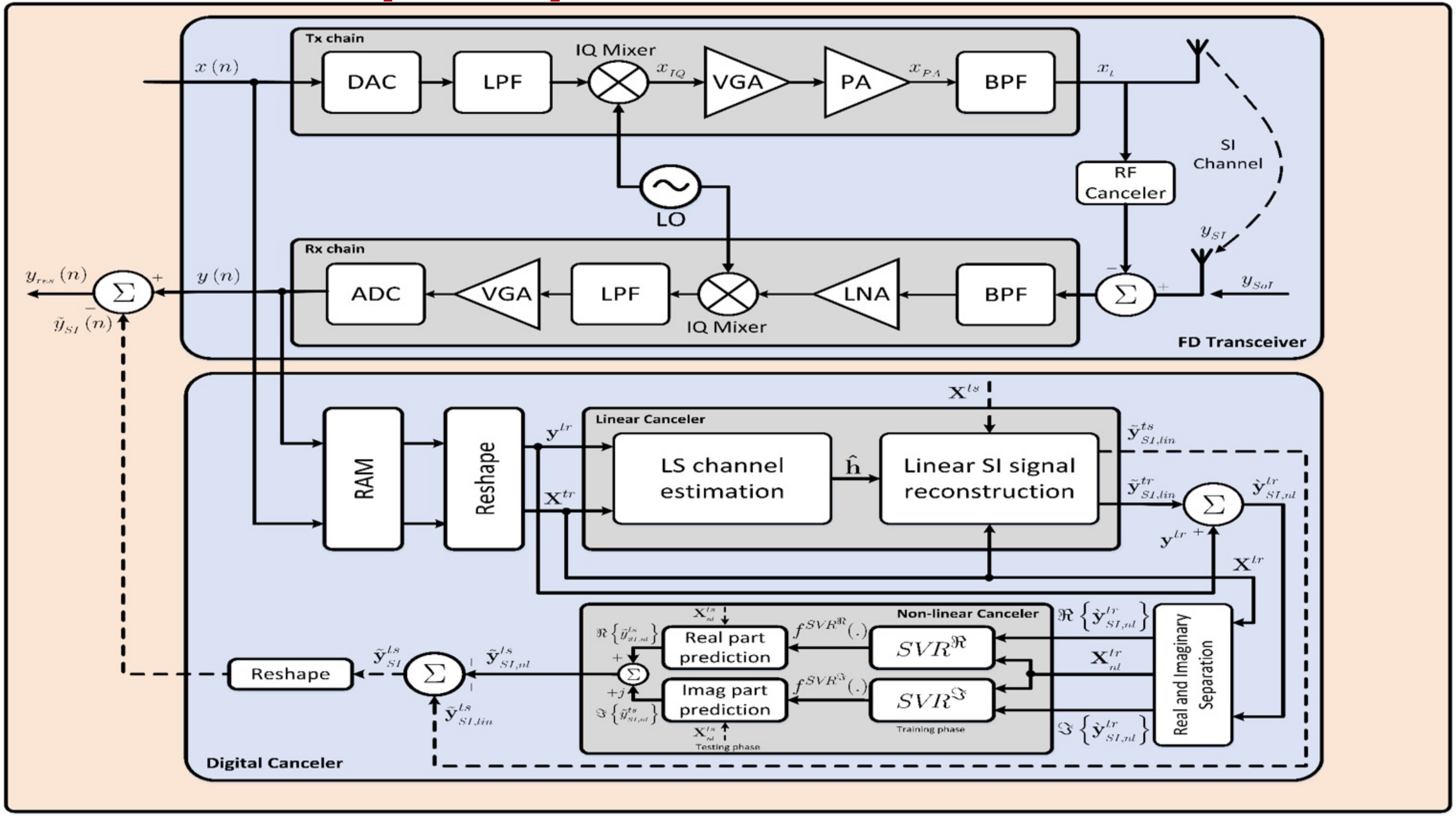
Slack variables are introduced to guarantee that the optimization problem is feasible



Linear and non-linear SVRs.

• A. Géron, Hands-on Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. Sebastopol, CA, USA: O'Reilly, 2017.

SVR-based full-duplex system model

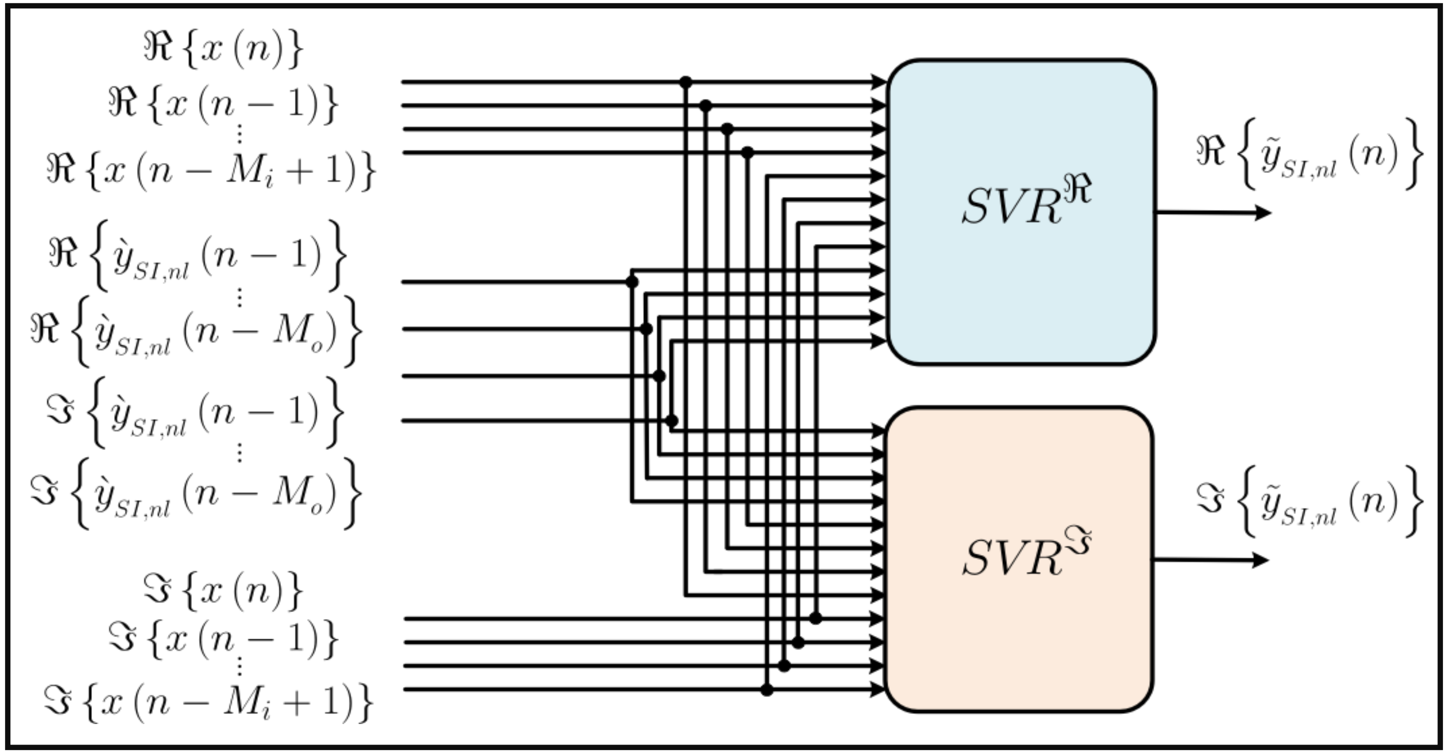


FD system model with linear and non-linear digital cancellation stages.

- A least-squares (LS)-based linear canceler: used to estimate the linear part of the SI signal
- SVR-based non-linear canceler: used to estimate the non-linear part of the SI signal

Proposed Output-feedback Time-delay SVR (OF-TDSVR)

- The input and output samples are utilized as features for training
- Specifically, the OF-TDSVR is fed by the real and imaginary parts of:
  - the current and past samples (M_i) of the input signal
  - past output samples (M_o) after applying the linear cancellation stage
- This can improve the learning capabilities of the OF-TDSVR and enhance its SIC compared to the SVR literature benchmarks



Proposed OF-TDSVR non-linear based canceler.

Note: Similar to the existing residual TDSVR (RTDSVR), the proposed OF-TDSVR also follows a residual scheme, where the non-linear cancellation is applied over the residual SI after performing the linear cancellation.

- M. Yilan, O. Gurbuz, and H. Ozkan, "Integrated linear and nonlinear digital cancellation for full duplex communication," *IEEE Wireless Communications*, Feb. 2021.

Achieved Results

- Datasets #1 and #2
- For a certain number of training sequence, find the peak performance (i.e., maximum SIC)
- Benchmarks: Tensor Completion & Deep Unfolding methods

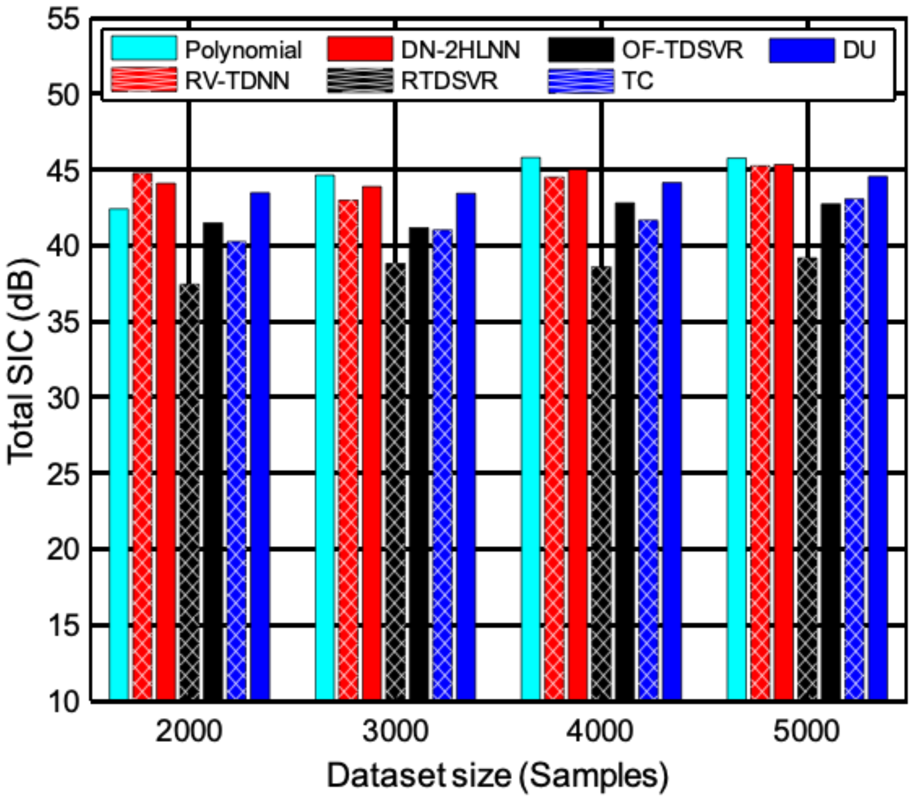
- F. Jochems and A. Balatsoukas-Stimming, "Non-linear self-interference cancellation via tensor completion," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, Nov. 2020, pp. 905–909.
- A. T. Kristensen, A. Burg, and A. Balatsoukas-Stimming, "Identification of non-linear RF systems using backpropagation," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, June 2020, pp. 1–6.

Achieved Results

MEMORIAL UNIVERSITY

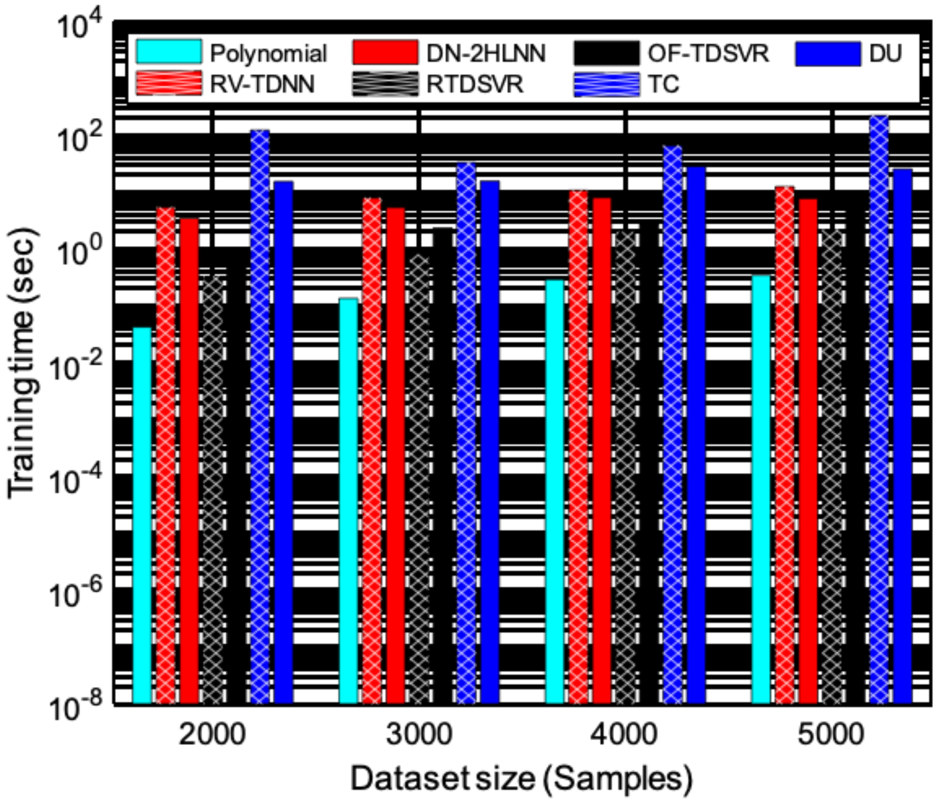
Public dataset #1

SIC Performance



Achieved SIC by different ML-based approaches compared to the polynomial-based canceler at various dataset sizes (public dataset #1).

Training time



Training time of different ML-based approaches compared to the polynomial-based canceler at various dataset sizes (public dataset #1).

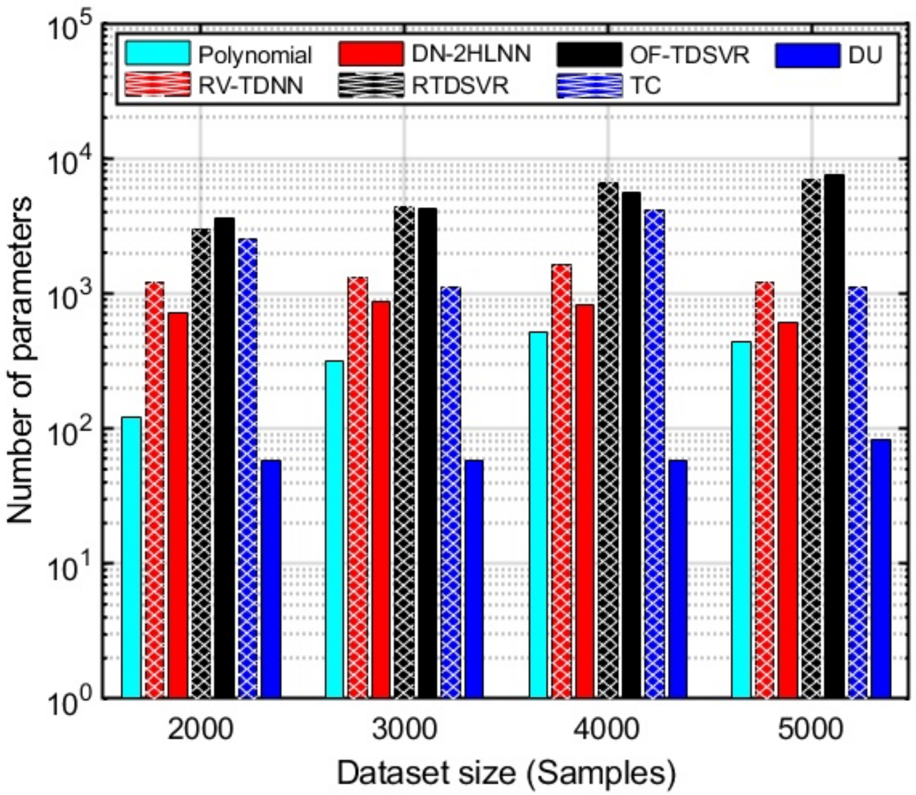
RV-TDNN: real-valued time delay NN; **DN-2HLNN:** dual-neurons two-hidden layers NN**; RTDSVR:** residual time-delay SVR; **OF-TDSVR**: output-feedback time-delay SVR; **TC:** tensor completion; **DU:** deep unfolding.

Achieved Results

MEMORIAL UNIVERSITY

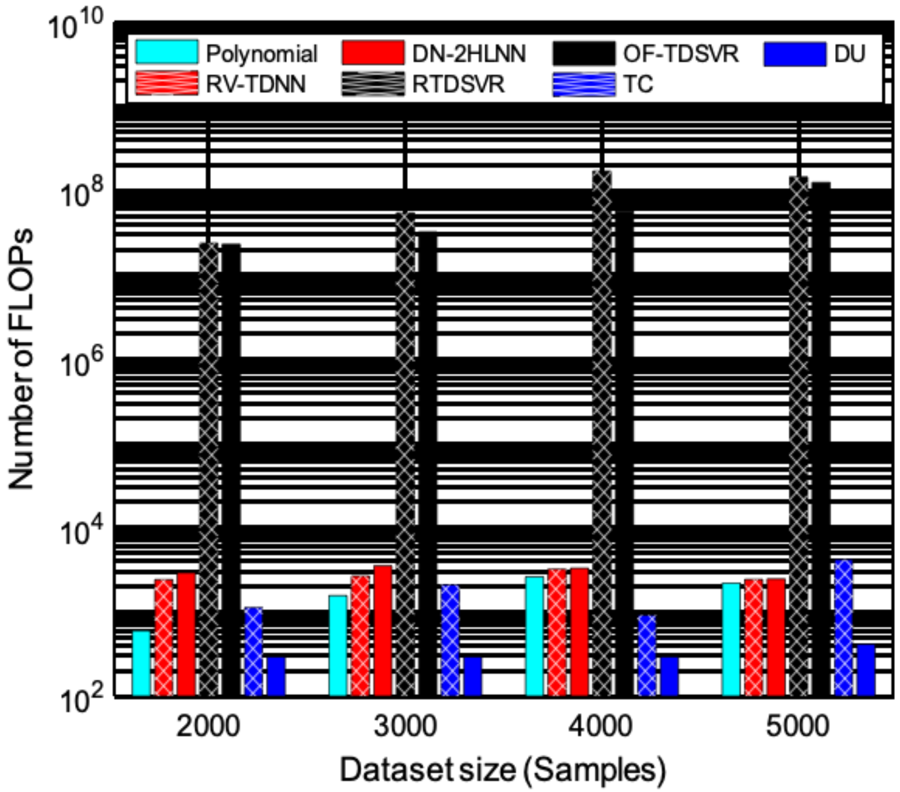
Public dataset #1

Memory storage



Total parameters of different ML-based approaches compared to the polynomial-based canceler at various dataset sizes (public dataset #1).

Computational complexity



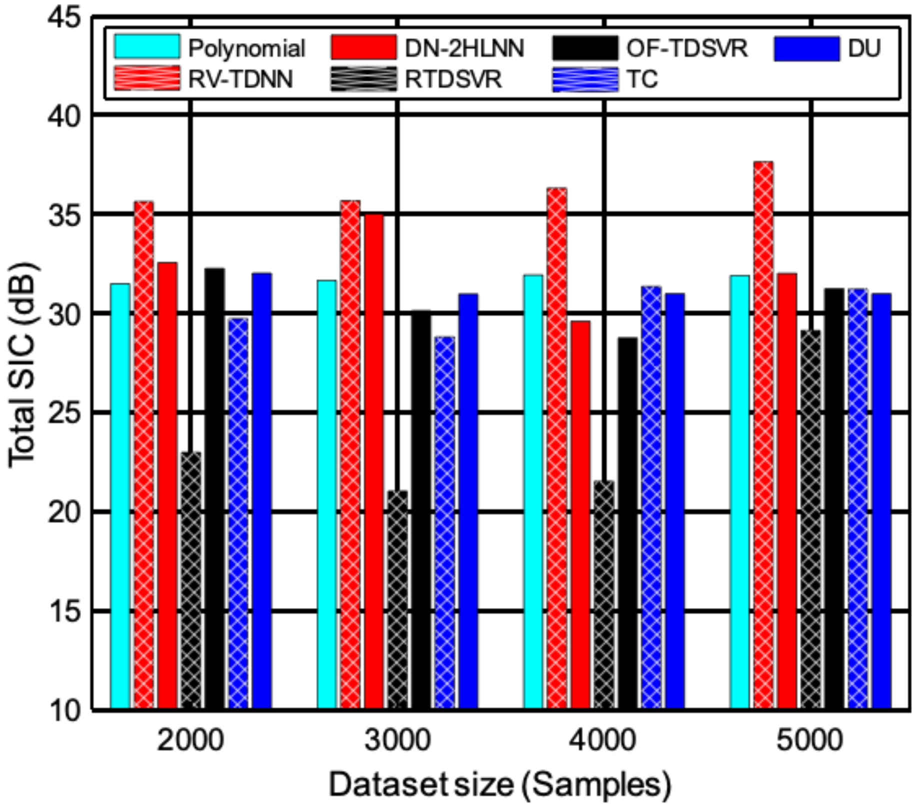
Total FLOPs of different ML-based approaches compared to the polynomial-based canceler at various dataset sizes (public dataset #1).

RV-TDNN: real-valued time delay NN; **DN-2HLNN:** dual-neurons two-hidden layers NN**; RTDSVR:** residual time-delay SVR; **OF-TDSVR**: output-feedback time-delay SVR; **TC:** tensor completion; **DU:** deep unfolding.

Achieved Results

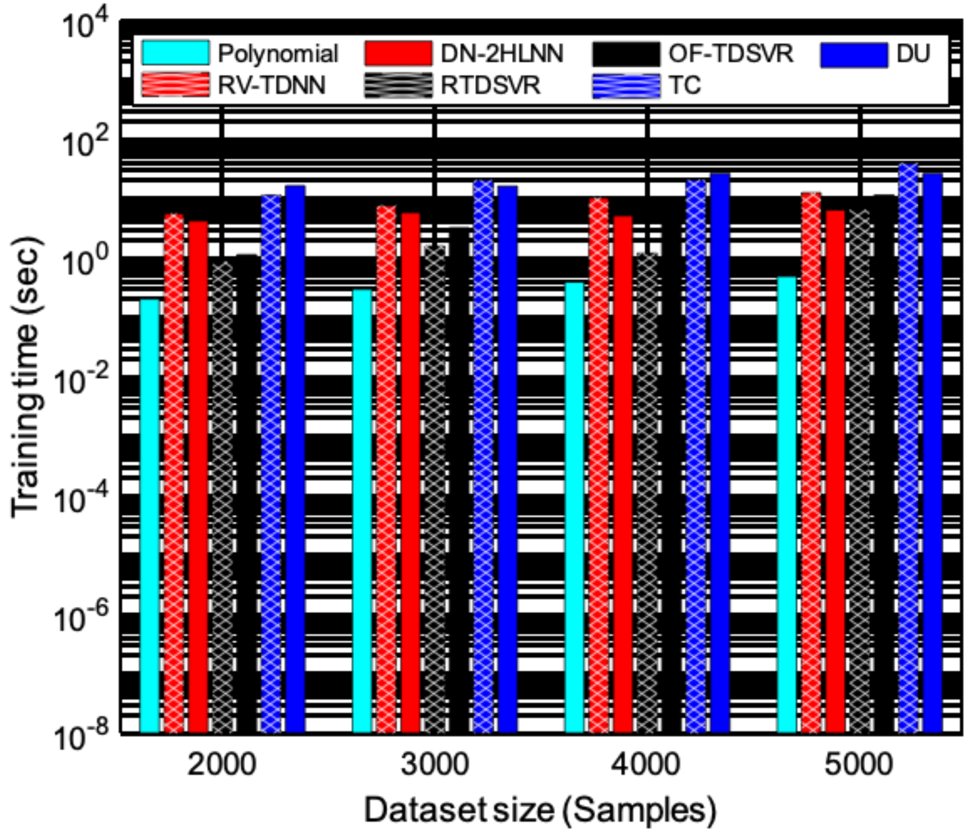
Public dataset #2

SIC Performance



Achieved SIC by different ML-based approaches compared to the polynomial-based canceler at various dataset sizes (public dataset #2).

Training time



Training time of different ML-based approaches compared to the polynomial-based canceler at various dataset sizes (public dataset #2).

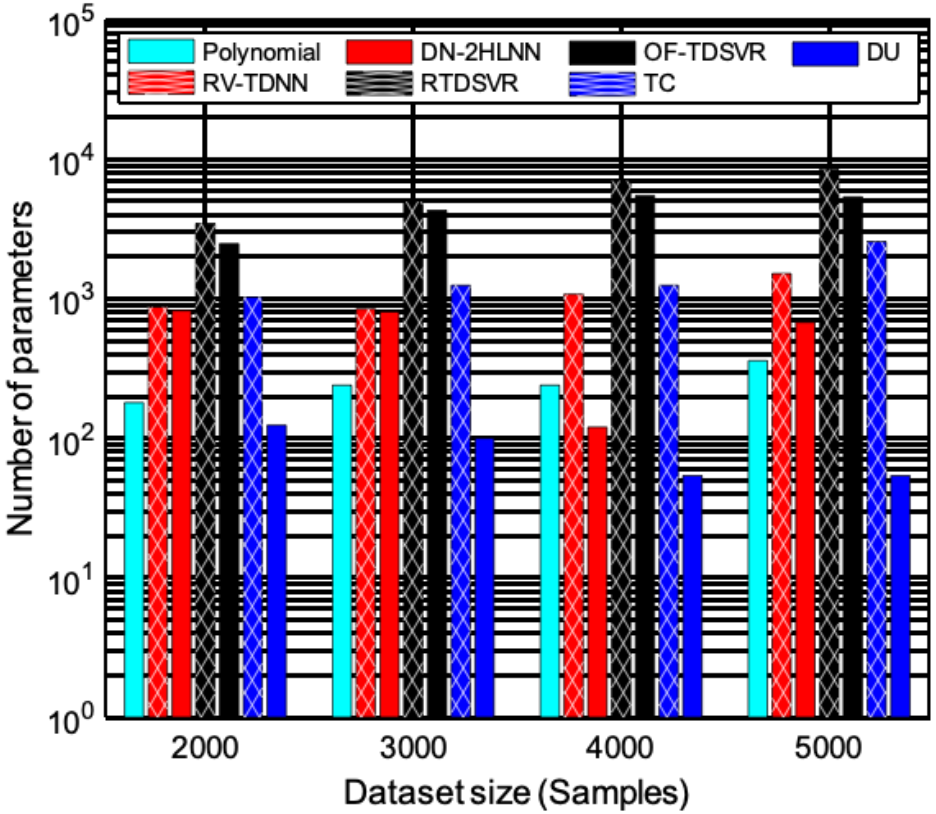
RV-TDNN: real-valued time delay NN; **DN-2HLNN:** dual-neurons two-hidden layers NN**; RTDSVR:** residual time-delay SVR; **OF-TDSVR**: output-feedback time-delay SVR; **TC:** tensor completion; **DU:** deep unfolding.

Achieved Results

MEMORIAL UNIVERSITY

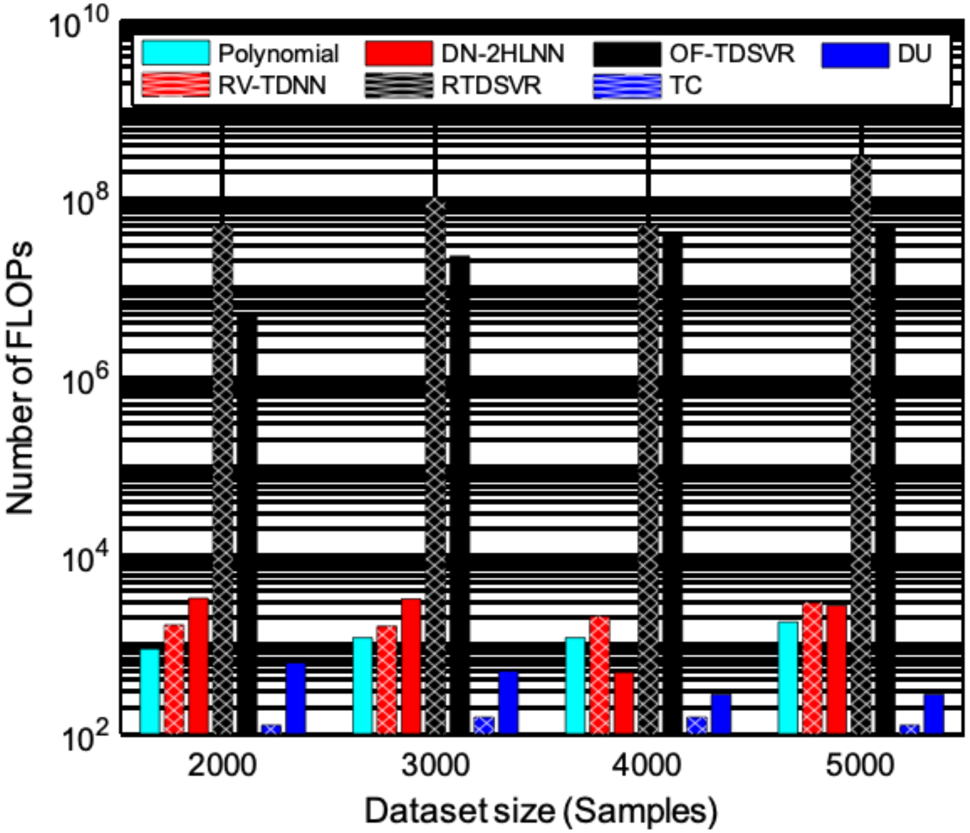
Public dataset #2

Memory storage



Total parameters of different ML-based approaches compared to the polynomial-based canceler at various dataset sizes (public dataset #2).

Computational complexity



Total FLOPs of different ML-based approaches compared to the polynomial-based canceler at various dataset sizes (public dataset #2).

RV-TDNN: real-valued time delay NN; **DN-2HLNN:** dual-neurons two-hidden layers NN**; RTDSVR:** residual time-delay SVR;
OF-TDSVR: output-feedback time-delay SVR; **TC:** tensor completion; **DU:** deep unfolding.

Performance Comparison - Canceler Efficiency

- Efficiency: η based on achieved SIC, fitting time, memory storage, and FLOPs

$$\eta = \frac{1}{w_c + w_\tau + w_\varrho + w_\mathcal{F}} \left(w_c \eta_c + w_\tau \eta_\tau + w_\varrho \eta_\varrho + w_\mathcal{F} \eta_\mathcal{F} \right)$$

$$\eta_c = \frac{C - C_{min}}{C_{max} - C_{min}}$$

$$\eta_\tau = 1 - \frac{\tau - \tau_{min}}{\tau_{max} - \tau_{min}}$$

$$\eta_\varrho = 1 - \frac{\varrho - \varrho_{min}}{\varrho_{max} - \varrho_{min}}$$

$$\eta_\mathcal{F} = 1 - \frac{\mathcal{F} - \mathcal{F}_{min}}{\mathcal{F}_{max} - \mathcal{F}_{min}}$$

- C : SIC achieved over a certain dataset with a particular average transmit power
- C_{max} C_{min}: maximum and minimum SIC
- τ: training time
- τ_{max} τ_{min}: maximum and minimum training time
- ϱ: number of parameters
- ϱ_{max} ϱ_{min} : maximum and minimum number of parameters
- $\mathcal{F}$: number of FLOPs
- $\mathcal{F}_{max}$ $\mathcal{F}_{min}$: maximum and minimum FLOPs

w_c w_τ w_ϱ and $w_\mathcal{F}$ ***are assigned to either 0 or 1 depending on the application requirements***

Achieved Results - Canceler Efficiency

Public dataset #1

Dataset Size	w_c	$w_{\mathcal{T}}$	w_{ϱ}	$w_{\mathcal{F}}$	Test case	η								
						Poly.	RV-TDNN	RNN	CV-TDNN	2HLNN	DN-2HLNN	OF-TDSVR	TC	DU
2000	1	0	0	0	SIC is the only system demand.		✓							
3000						✓								
4000						✓								
5000						✓								
2000	1	1	0	0	SIC and training time are the only system demands.		✓							
3000						✓								
4000						✓								
5000						✓								
2000	1	0	1	0	SIC and memory are the only system demands.									✓
3000						✓								
4000						✓								
5000						✓								
2000	1	0	0	1	SIC and complexity are the only system demands.		✓							
3000						✓								
4000						✓								
5000						✓								

Dataset Size	w_c	$w_{\mathcal{T}}$	w_{ϱ}	$w_{\mathcal{F}}$	Test case	η								
						Poly.	RV-TDNN	RNN	CV-TDNN	2HLNN	DN-2HLNN	OF-TDSVR	TC	DU
2000	1	1	1	0	SIC, training time, and memory are the only system demands.					✓				
3000						✓								
4000						✓								
5000						✓								
2000	1	1	0	1	SIC, training time, and complexity are the only system demands.		✓							
3000						✓								
4000						✓								
5000						✓								
2000	1	0	1	1	SIC, memory, and complexity are the only system demands.					✓				
3000						✓								
4000						✓								
5000						✓								
2000	1	1	1	1	SIC, training time, memory, and complexity are all system demands.					✓				
3000						✓								
4000						✓								
5000						✓								

RV-TDNN: real-valued time delay NN; **RNN:** recurrent NN; **CV-TDNN:** complex-valued time-delay NN; **2HLNN:** two-hidden layers NN;
DN-2HLNN: dual-neurons two-hidden layers NN; **RTDSVR:** residual time-delay SVR; OF-TDSVR: output-feedback time-delay SVR; **TC:** tensor completion; **DU:** deep unfolding.

Achieved Results - Canceler Efficiency

MEMORIAL UNIVERSITY

Public dataset #2

Dataset Size	w_c	$w_{\mathcal{T}}$	w_{ϱ}	$w_{\mathcal{F}}$	Test case	η: Poly.	RV-TDNN	RNN	CV-TDNN	2HLNN	DN-2HLNN	OF-TDSVR	TC	DU
2000	1	0	0	0	SIC is the only system demand.		✓							
3000							✓							
4000							✓							
5000							✓							
2000	1	1	0	0	SIC and training time are the only system demands.							✓		
3000						✓								
4000						✓								
5000							✓							
2000	1	0	1	0	SIC and memory are the only system demands.		✓							
3000							✓							
4000							✓							
5000							✓							
2000	1	0	0	1	SIC and complexity are the only system demands.		✓							
3000							✓							
4000							✓							
5000							✓							

Dataset Size	w_c	$w_{\mathcal{T}}$	w_{ϱ}	$w_{\mathcal{F}}$	Test case	η: Poly.	RV-TDNN	RNN	CV-TDNN	2HLNN	DN-2HLNN	OF-TDSVR	TC	DU
2000	1	1	1	0	SIC, training time, and memory are the only system demands.	✓								
3000						✓								
4000						✓								
5000							✓							
2000	1	1	0	1	SIC, training time, and complexity are the only system demands.	✓								
3000						✓								
4000						✓								
5000							✓							
2000	1	0	1	1	SIC, memory, and complexity are the only system demands.		✓							
3000							✓							
4000							✓							
5000							✓							
2000	1	1	1	1	SIC, training time, memory, and complexity are all system demands.	✓								
3000						✓								
4000						✓								
5000							✓							

RV-TDNN: real-valued time delay NN; **RNN:** recurrent NN; **CV-TDNN:** complex-valued time-delay NN; **2HLNN:** two-hidden layers NN; **DN-2HLNN:** dual-neurons two-hidden layers NN; **RTDSVR:** residual time-delay SVR; OF-TDSVR: output-feedback time-delay SVR; **TC:** tensor completion; **DU:** deep unfolding.

Summary (reduced data set for training)

Conclusion for dataset #1

Promising Solutions:

- **Polynomial-based canceler:** highest SIC with the **lowest training time**
- **DU-based canceler:** requires the **lowest number of parameters** and **FLOPs**, albeit at the cost of **reduced SIC** and **increased training time**

Conclusion for dataset #2

Promising Solutions:

- **RV-TDNN-based canceler:** **highest SIC** with reasonable memory storage and computational complexity
- **DU/TU-based canceler:** requires the **lowest number of parameters/FLOPs**, albeit at the cost of reduced SIC and increased training time

Conclusion

Conclusion: full data set used for training

NN-based SIC: Model-Centric Approach

- **DN-2HLNN**: lowest complexity and provides about 60% reduction in the number of network parameters and FLOPs over the polynomial-based canceler
- When trained for dataset #1, it performs reasonably well when applied to dataset #2

Conclusion: reduced data set used for training

Model-Centric Approach

- **Higher power level**: **RV-TDNN**-based canceler: highest SIC with reasonable memory storage and computational complexity
- **Lower power level**: polynomial canceler: highest SIC with the lowest training time

Ongoing and Future Work

MEMORIAL UNIVERSITY

Machine Learning-based Self-interference Cancellation

- **Train and test solutions for various parameters:** modulation formats, powers, sampling frequencies
- **Follow a data-centric approach:** input/output data samples are captured before/after the DAC/ADC
- **Online learning and extreme learning machine:** performance-complexity-training time to adapt to changes
- **Generalization**: out of distrbution generalization/model generalization
- **Study when using the signal-of-interest (SoI)** & what if the SoI uses frequency domain modulation formats?
- **Extension to multiple-input multiple-output (MIMO) case:** complexity linearly increases under MIMO operation

Full Duplex Communications – Network

- **Interference-limited scenario:** beamforming, scheduling, multiple access, resource allocation
- **Integrated sensing and communication (ISAC):** channels estimation, interference-limited communications