Binary MIMO Detection via Extreme Learning Machines in 5G Network and Beyond

Wai-Yiu Keung[†,§] and Umair M. Qureshi[†]

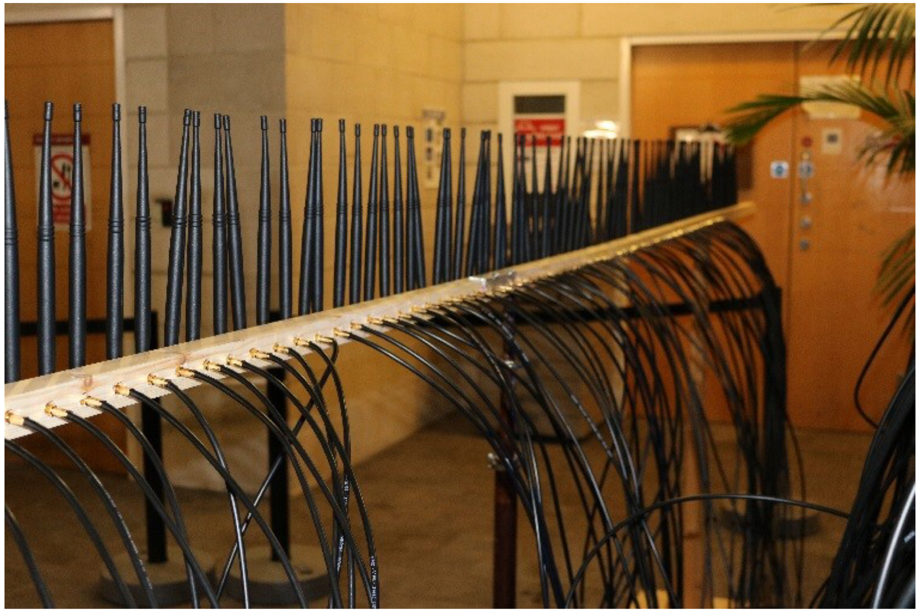
[†]Dept. of Comp. Sci. & Eng., [§]Dept. of Elec. Eng.

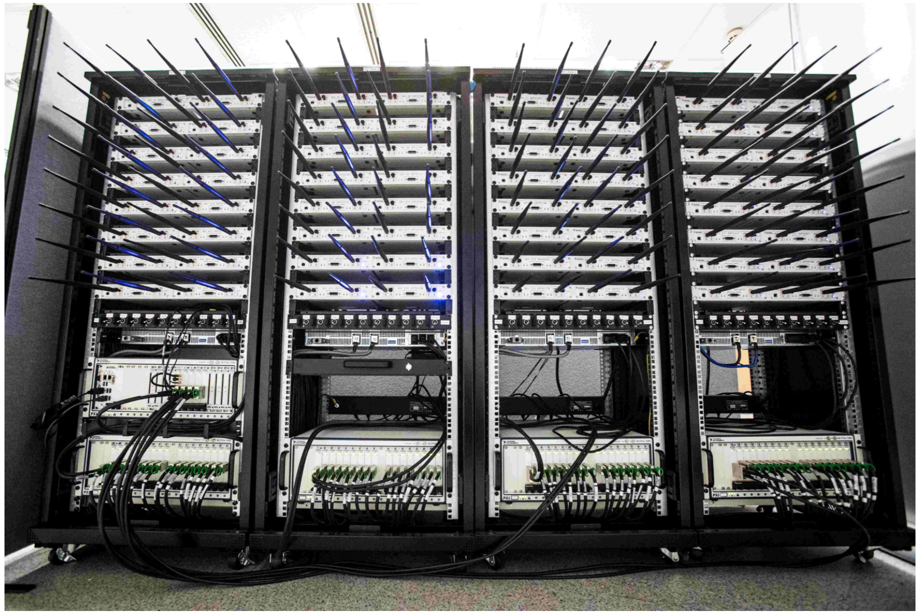
香港中文大學
The Chinese University of Hong Kong

SRIN Workshop, IEEE Globecom, December 8, 2023

Contact: `wykeung`, `umair [at] cse.cuhk.edu.hk`

Motivation

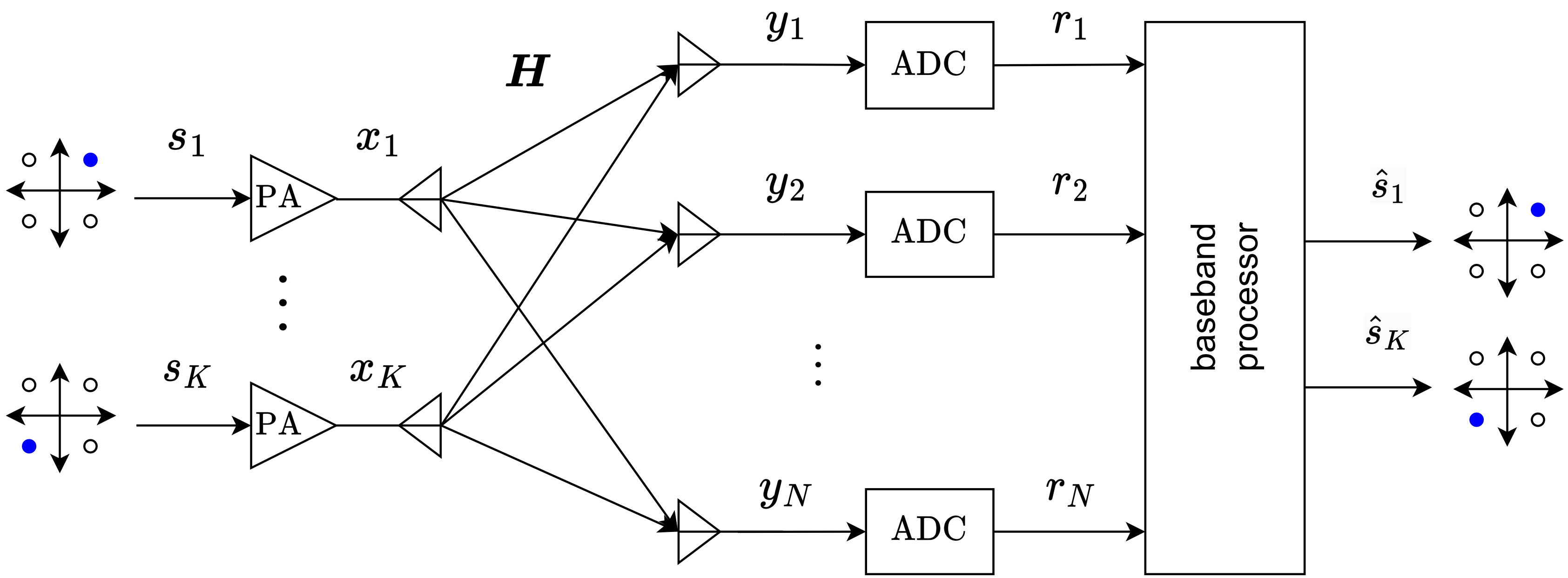




Source: P. Harris et al., "*An overview of massive MIMO research at the University of Bristol*," Radio Propagation and Technologies for 5G (2016), Durham, UK, 2016, pp. 1-5.

- MIMO is expected to scale large in future gen. of comm. sys.
- premise of spatial mux. benefits in faster tx. rate, wider coverage, better QoS
- comes at a price: increasing hardware complexity and cost

Classical MIMO Up-link

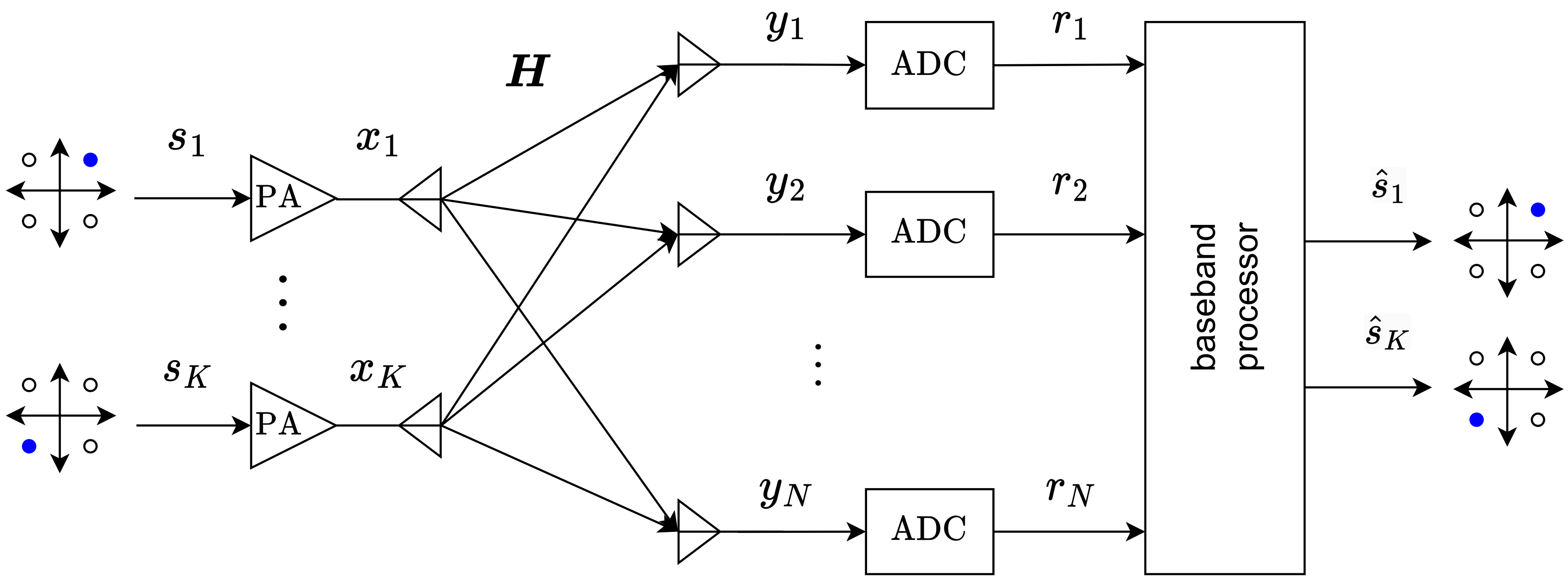


- with ideal assumptions on PA/DACs, up-link signal model:

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \text{noise}$$

where $\boldsymbol{y} \in \mathbb{C}^N$ is the rx. signal at the BS; $\boldsymbol{H}$ is the channel response; $\boldsymbol{x} = \boldsymbol{s} \in \mathcal{S}^K$ are the information-carrying signals shot by the users

Classical MIMO Up-link



- with ideal assumptions on PA/DACs, up-link signal model:

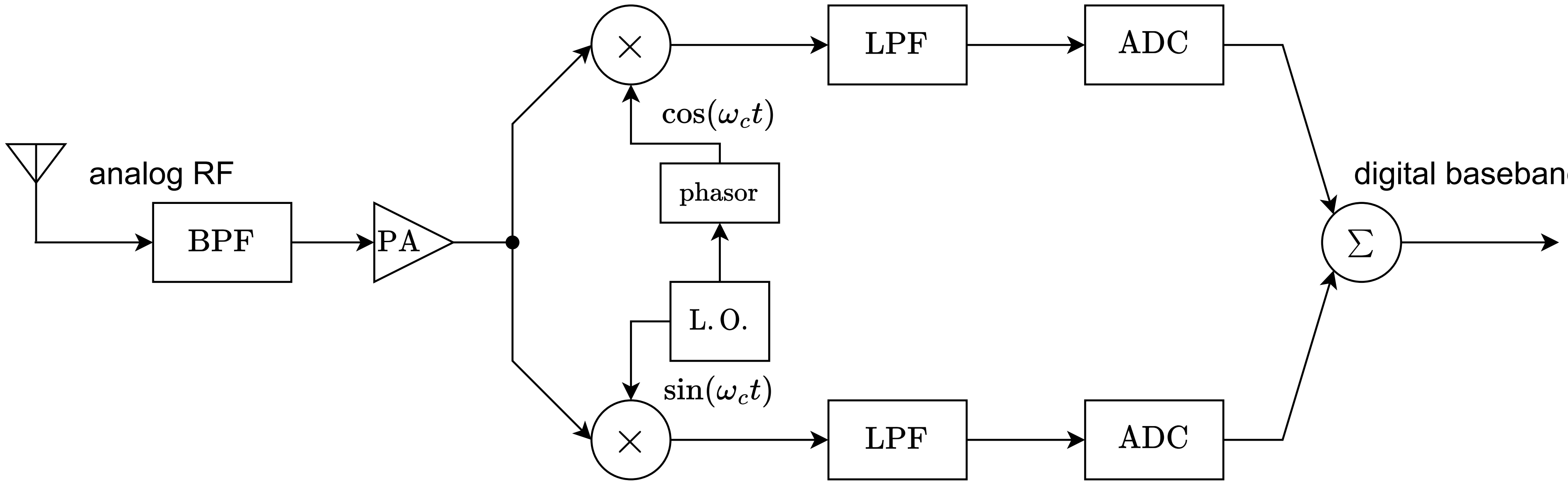
$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \text{noise}$$

  where $\boldsymbol{y} \in \mathbb{C}^N$ is the rx. signal at the BS; $\boldsymbol{H}$ is the channel response; $\boldsymbol{x} = \boldsymbol{s} \in \mathcal{S}^K$ are the information-carrying signals shot by the users

- detection: given the channel state information $\boldsymbol{H}$, estimate $\boldsymbol{s}$ from the noisy observation vector $\boldsymbol{y}$

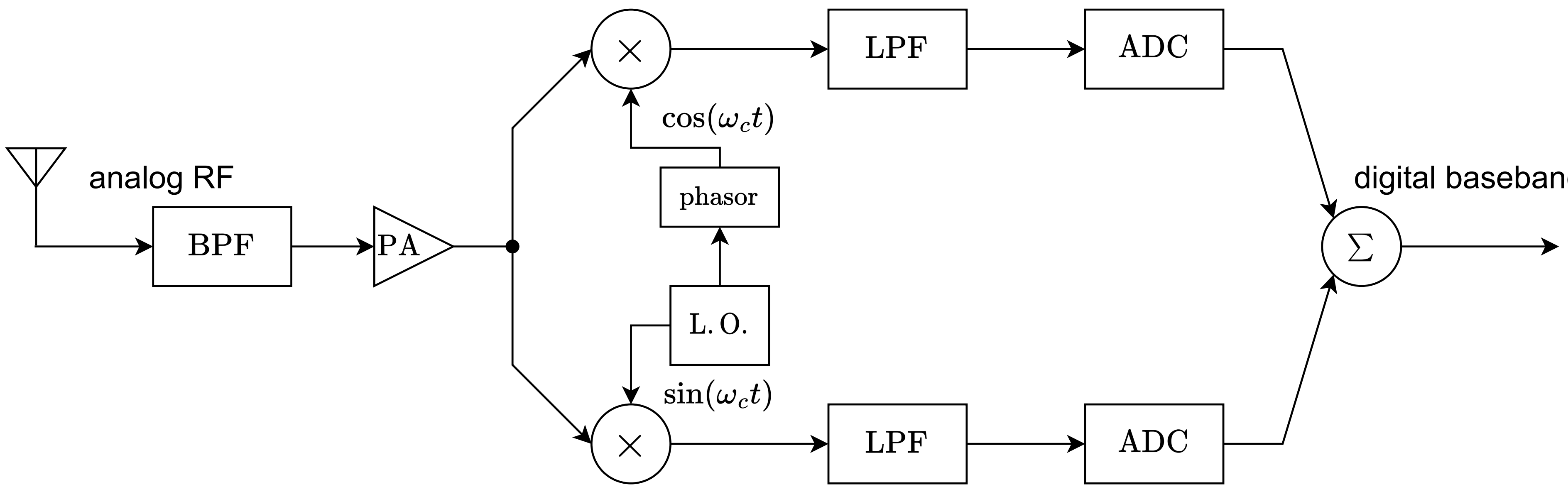
- classical problem, well studied in the literature; but...

BS Side: Up-link Receiver with Cheap ADCs



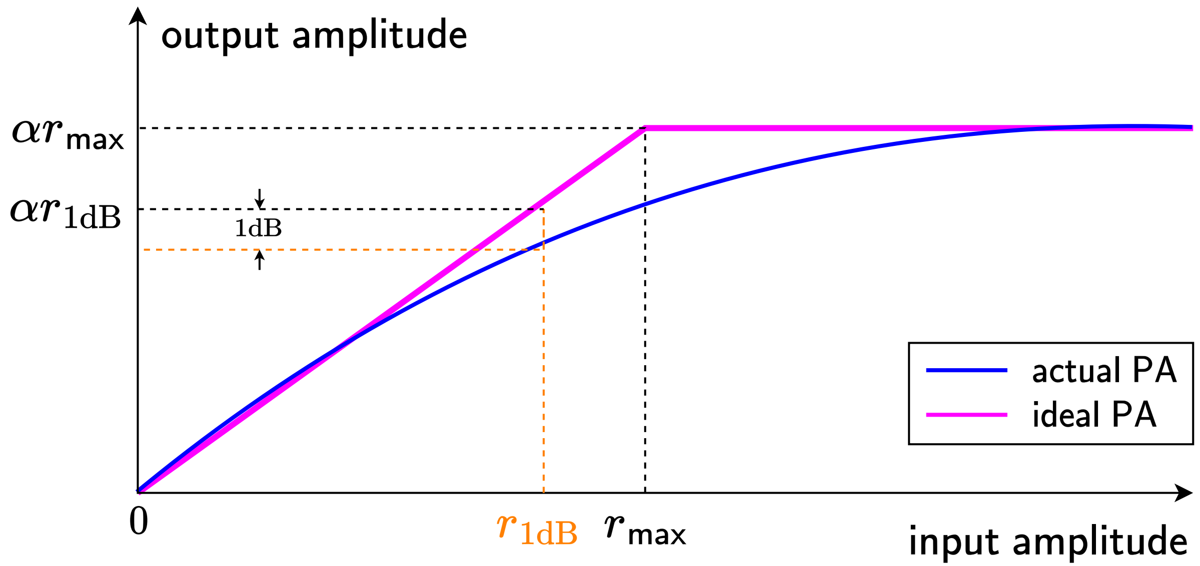
- high res. ADC marks the major power burden in rx. implementation

BS Side: Up-link Receiver with Cheap ADCs



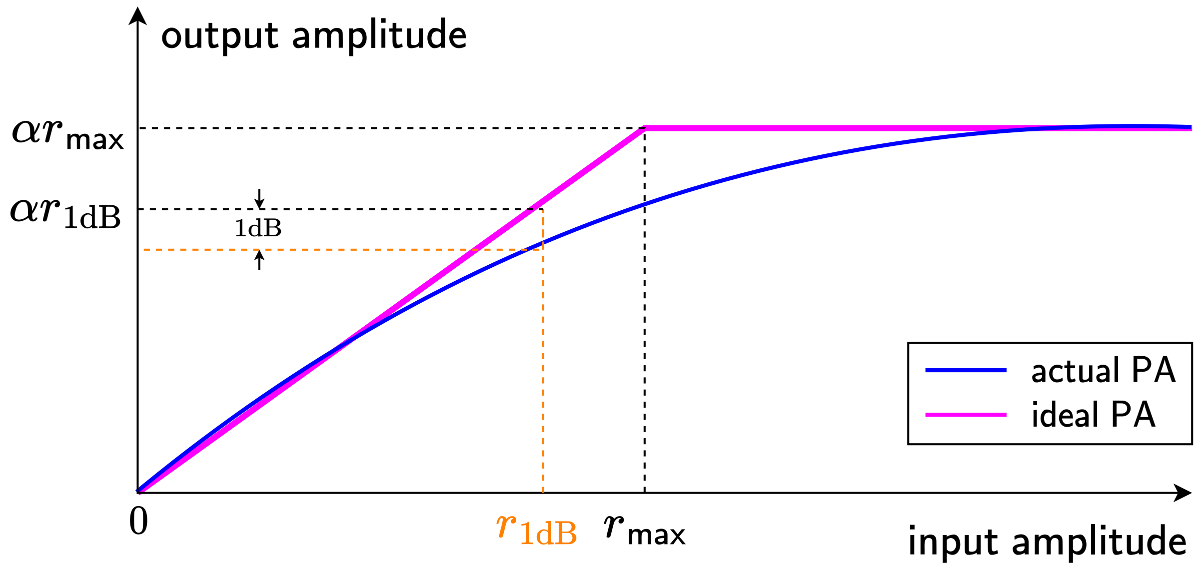
- high res. ADC marks the major power burden in rx. implementation
- recent trend is to replace them with cheaper converters, e.g. one-bit ADCs; allows cheaper PAs at the rx. implementation too
- **setback:** heavy quantization error...

UE Side: Non-linear PA Effect

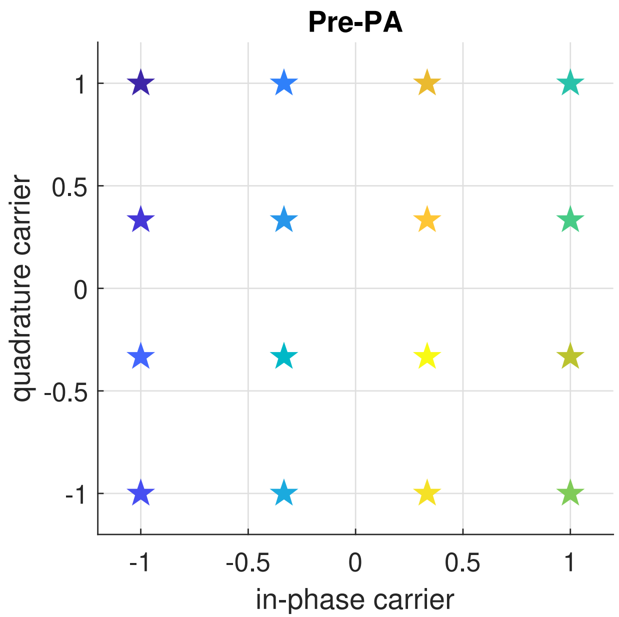


- non-linear PAs may destroy the signal constellation structure

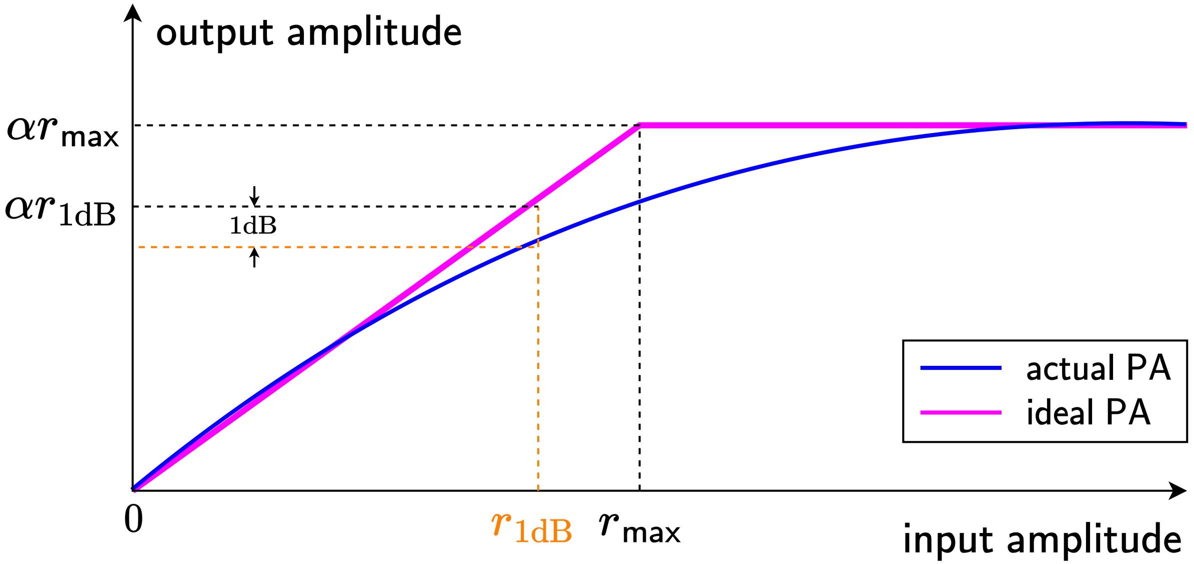
UE Side: Non-linear PA Effect



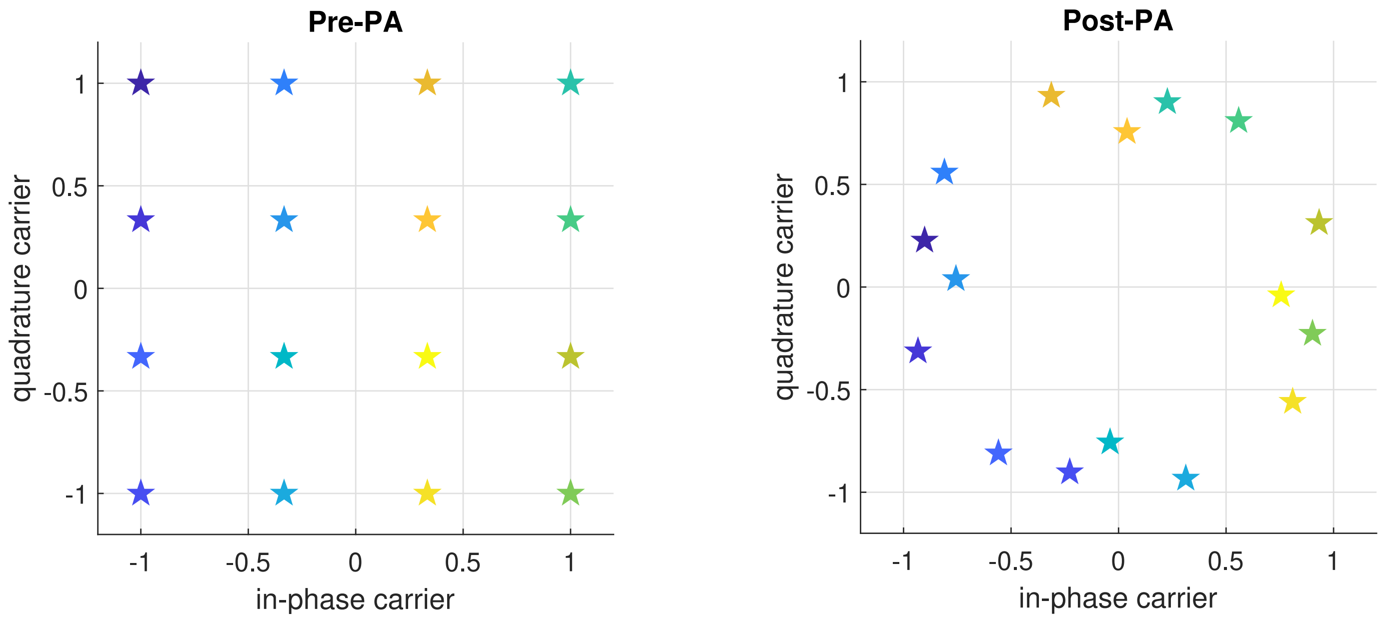
- non-linear PAs may destroy the signal constellation structure



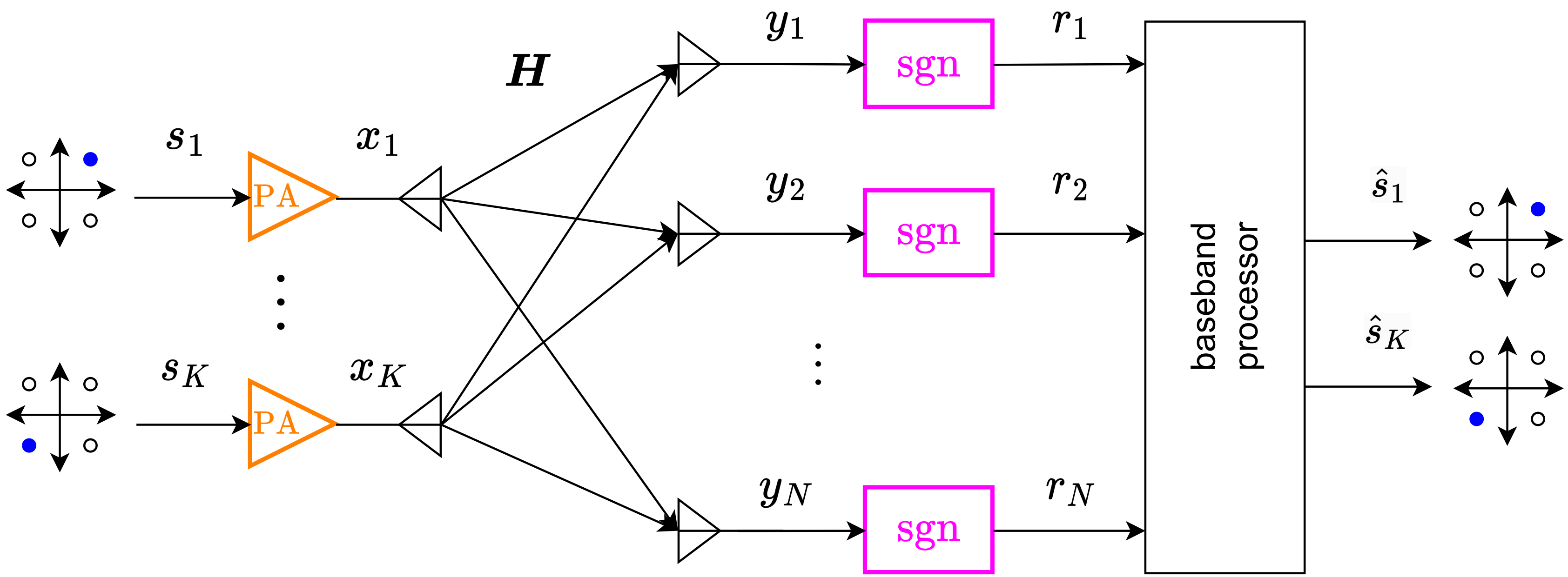
UE Side: Non-linear PA Effect



- non-linear PAs may destroy the signal constellation structure

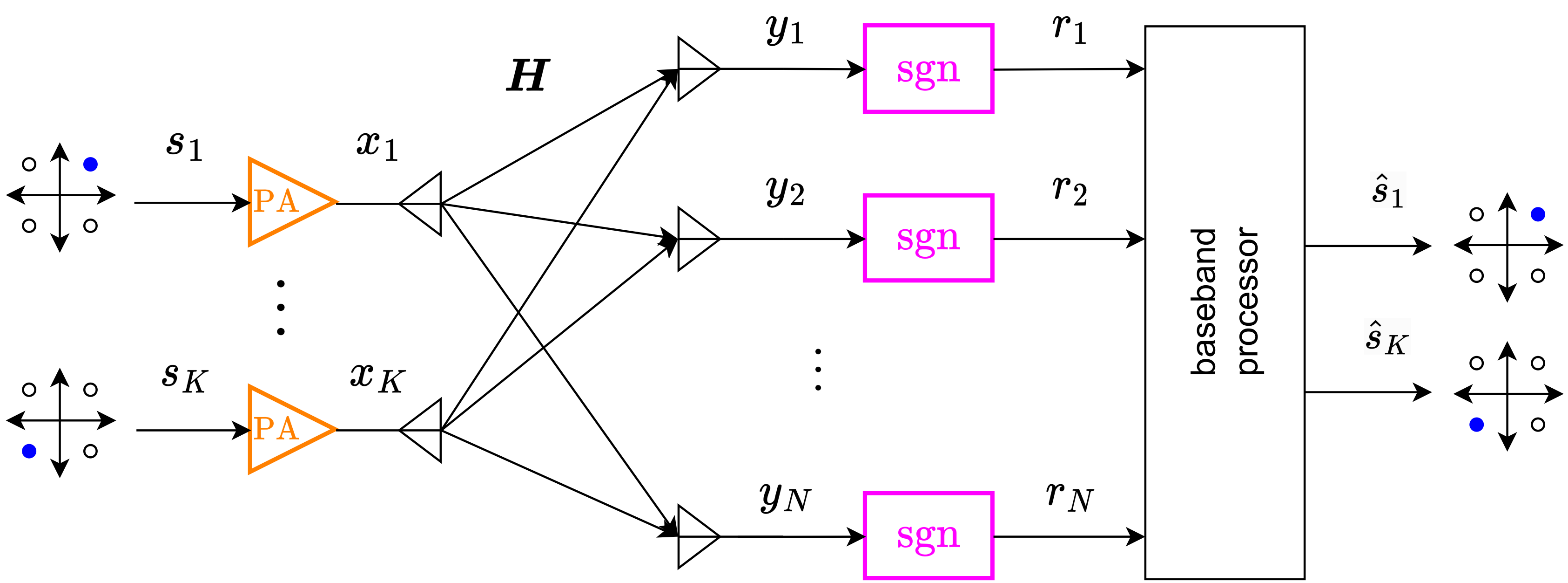


Binary MIMO Up-link with PA effects at UE



- signal model with **one-bit** DACs at BS and non-ideal PA at UE: $\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \text{noise}$;
  - $\boldsymbol{r} = \text{sgn}(\boldsymbol{y}) \in \{\pm 1, \pm \mathrm{j}\}^N$ and $\boldsymbol{x} = \Psi(\boldsymbol{s}) \in \mathbb{C}^K$; $\Psi(\cdot)$ is the PA effect

Binary MIMO Up-link with PA effects at UE



- signal model with **one-bit** DACs at BS and non-ideal PA at UE: $\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \text{noise}$;
  - $\boldsymbol{r} = \mathrm{sgn}(\boldsymbol{y}) \in \{\pm 1, \pm \mathrm{j}\}^N$ and $\boldsymbol{x} = \Psi(\boldsymbol{s}) \in \mathbb{C}^K$; $\Psi(\cdot)$ is the PA effect

- **challenge:** conventional detectors (such as ZF, SDR, sphere decoding etc.) do not work because the problem structure is destroyed by PA/DACs

- **recent trend:** apply machine learning/ deep learning/ deep unfolding

Existing Works

- conventional SP-comm. methods to handle the quantization error at DACs **[MKN07, RPL14, CMH16]**
- use model-driven machine learning for MIMO detection **[SDW19, HWJL20, NNT$^+$23]**
- deep unfolding — which is to build a DNN with inspirations taken from an iterative algo. customised for MIMO detection **[MLE21, NSN21, SM21]**

Extreme Learning Machines: A Heuristic Solver

- idea: consider the functional approximater, with $g(\cdot)$ being an activation:

$$f(\boldsymbol{r}; \boldsymbol{\Theta}) = g(\boldsymbol{W}\boldsymbol{r} + \boldsymbol{b})\ \boldsymbol{A}$$

where $\boldsymbol{\Theta} = (\boldsymbol{W}, \boldsymbol{b})$ are randomly initialized (fixed); the only var. for opt. is $\boldsymbol{A}$

Extreme Learning Machines: A Heuristic Solver

- idea: consider the functional approximater, with $g(\cdot)$ being an activation:

$$f(\boldsymbol{r};\boldsymbol{\Theta}) = g(\boldsymbol{W}\boldsymbol{r}+\boldsymbol{b})\ \boldsymbol{A}$$

where $\boldsymbol{\Theta} = (\boldsymbol{W}, \boldsymbol{b})$ are randomly initialized (fixed); the only var. for opt. is $\boldsymbol{A}$

- learning target: given the pilot symbols and the received signals $\{\boldsymbol{s}^{(i)}, \boldsymbol{r}^{(i)}\}_{i=1}^{m}$

$$\text{find} \quad \boldsymbol{A} \quad \text{s.t.} \quad f(\boldsymbol{r}^{(i)};\boldsymbol{\Theta}) \approx \boldsymbol{s}^{(i)}\ \forall i = 1, \ldots, m$$

which is easily obtained by the Moore-Penrose inverse

Extreme Learning Machines: A Heuristic Solver

- idea: consider the functional approximater, with $g(\cdot)$ being an activation:

$$f(\boldsymbol{r}; \boldsymbol{\Theta}) = g(\boldsymbol{W}\boldsymbol{r} + \boldsymbol{b})\ \boldsymbol{A}$$

  where $\boldsymbol{\Theta} = (\boldsymbol{W}, \boldsymbol{b})$ are randomly initialized (fixed); the only var. for opt. is $\boldsymbol{A}$

- learning target: given the pilot symbols and the received signals $\{\boldsymbol{s}^{(i)}, \boldsymbol{r}^{(i)}\}_{i=1}^{m}$

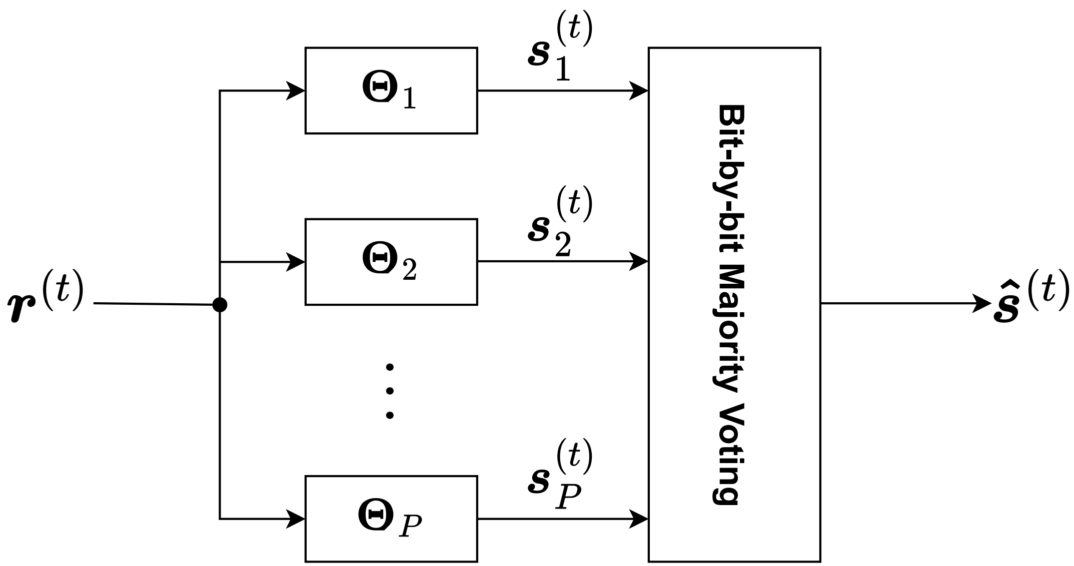
$$\text{find} \quad \boldsymbol{A} \quad \text{s.t.} \quad f(\boldsymbol{r}^{(i)}; \boldsymbol{\Theta}) \approx \boldsymbol{s}^{(i)}\ \forall i = 1, \ldots, m$$

  which is easily obtained by the Moore-Penrose inverse

- an old idea in functional approx.; dates back to 1995 under the name of random vector functional-link network **[IP95]**

- ELM has been applied to large MIMO up-link, e.g. **[GGE21, CDCC21, CCD+22]**; but it is the first time for the one-bit case

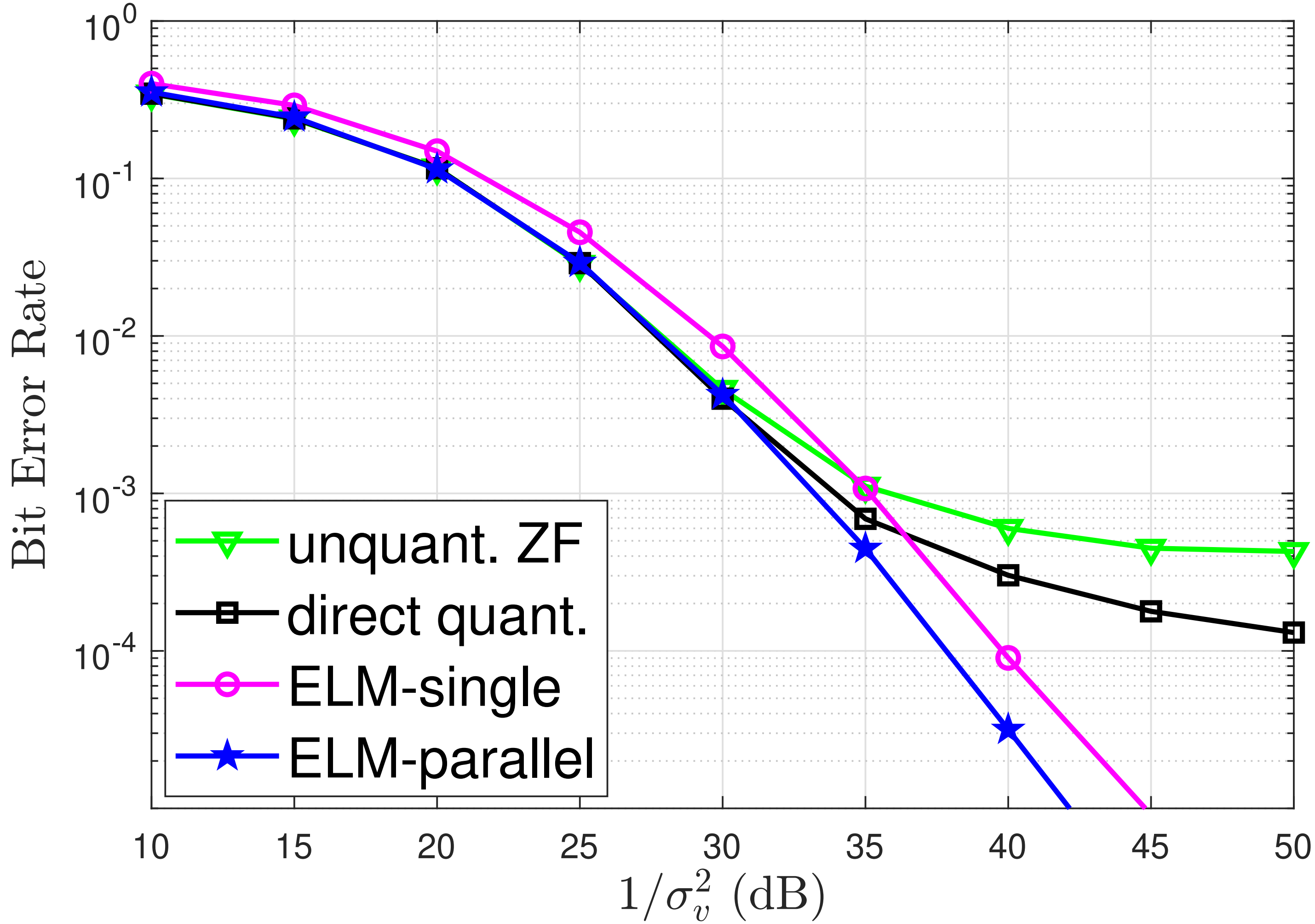
Parallel ELM Detector — Yet Another Heuristic



- since the init. $\boldsymbol{\Theta}$ is random, it is expected that the resultant f is not optimum
- our heuristic attempt: initialize P parallel $\boldsymbol{\Theta}_p$'s and train $\boldsymbol{A}_p$'s individually, and perform majority voting on each element of $\boldsymbol{s}$

$$\hat{\boldsymbol{s}} = \mathrm{dec}\left(\sum_{p=1}^{P} f(\boldsymbol{r}, \boldsymbol{A}_p; \boldsymbol{\Theta}_p)\right)$$

Simulation Result: Bit Error Rate



- settings: $N = 30$ rx. antenna at BS; $K = 8$ tx. UEs; 2048 hidden units; $P = 7$ parallel ELMs;
- the benchmark of ZF (both quant. and unquant.) is performed by estimating $\boldsymbol{H}_{\rm est}$ using the full res. rx. signals

Parameter Reduction by Circulant Matrix Initialization

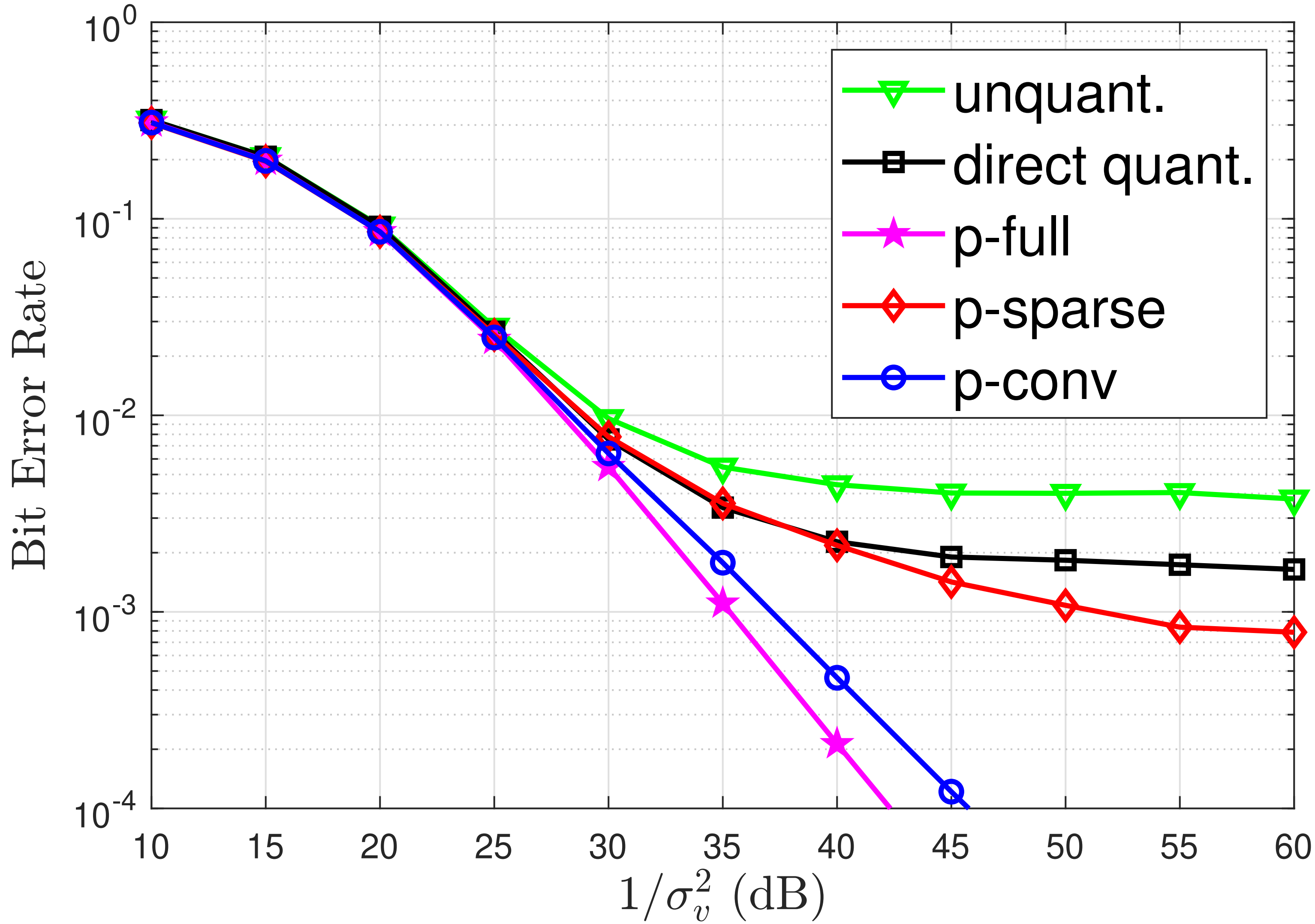
$$\boldsymbol{W} = \begin{bmatrix} w_1 & w_2 & \cdots & w_V & 0 & 0 & \cdots & 0 \\ 0 & w_1 & w_2 & \cdots & w_V & 0 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \ddots & & \vdots \\ 0 & 0 & \cdots & 0 & w_1 & w_2 & \cdots & w_V \end{bmatrix}$$

- reduce parameters needed to be stored in $f(\cdot, \boldsymbol{\Theta}, \boldsymbol{A})$ by using this structure

Units	Metrics	Full	Sparse	Conv.
128	Parameters	15.35	3.67	3.68
	Training Error	**0.1462%**	0.3285%	**0.1494%**
512	Parameters	58.62	11.94	12.25
	Training Error	**0.0246%**	0.3358%	**0.0400%**
2048	Parameters	231.73	45	46.52
	Training Error	**0.0027%**	0.2932%	**0.0027%**

Table: Parameters needed (in million) and the training error of different init

Simulation Result: Bit Error Rate



$N = 20, K = 4, P = 17, 1024$ hidden units

Take-home Points

- large BS appears to be the trend for future generation of comm. sys.
- using one-bit ADCs at the BS helps to improve power efficiency for the overall sys
- we proposed to apply ELM in the context of one-bit MIMO detection, offering a computationally friendly approach for the challenging task

Take-home Points

- large BS appears to be the trend for future generation of comm. sys.
- using one-bit ADCs at the BS helps to improve power efficiency for the overall sys
- we proposed to apply ELM in the context of one-bit MIMO detection, offering a computationally friendly approach for the challenging task

Thank you!

Key References I

[CCD+22] Diego-Fernando Carrera, Vargas-Rosales Cesar, Zabala-Blanco David, et al., *Novel multilayer extreme learning machine as a massive MIMO receiver for millimeter wave communications*, IEEE Access **10** (2022), 58965–58981.

[CDCC21] Diego-Fernando Carrera, Zabala-Blanco David, Vargas-Rosales Cesar, and Azurdia-Meza Cesar, *Extreme learning machine-based receiver for multi-user massive MIMO systems*, IEEE Comm. Letters **25** (2021), no. 2, 484–488.

[CMH16] Junil Choi, Jianhua Mo, and Robert W. Heath, *Near maximum-likelihood detector and channel estimator for uplink multiuser massive MIMO systems with one-bit ADCs*, IEEE Trans. on Comm. **64** (2016), no. 5, 2005–2018.

[GGE21] Dawei Gao, Qinghua Guo, and Yonina C. Eldar, *Massive MIMO as an extreme learning machine*, IEEE Trans. on Veh. Tech. **70** (2021), no. 1, 1046–1050.

[HWJL20] Hengtao He, Chao-Kai Wen, Shi Jin, and Geoffrey Ye Li, *Model-driven deep learning for MIMO detection*, IEEE Trans. on Sig. Proc. **68** (2020), 1702–1715.

[IP95] B. Igelnik and Yoh-Han Pao, *Stochastic choice of basis functions in adaptive function approximation and the functional-link net*, IEEE Trans. on Neural Networks **6** (1995), no. 6, 1320–1329.

Key References II

[MKN07] Amine Mezghani, Mohamed-Seifeddine Khoufi, and Josef Nossek, *A modified MSE receiver for quantized MIMO systems*.

[MLE21] Vishal Monga, Yuelong Li, and Yonina C. Eldar, *Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing*, IEEE Sig. Proc. Mag. **38** (2021), no. 2, 18–44.

[NNT$^+$23] Ly V. Nguyen, Nhan T. Nguyen, Nghi H. Tran, Markku Juntti, A. Lee Swindlehurst, and Duy H. N. Nguyen, *Leveraging deep neural networks for massive MIMO data detection*, IEEE Wireless Comm. **30** (2023), no. 1, 174–180.

[NSN21] Ly V. Nguyen, A. Lee Swindlehurst, and Duy H. N. Nguyen, *Linear and deep neural network-based receivers for massive MIMO systems with one-bit ADCs*, IEEE Trans. on Wireless Comm. **20** (2021), no. 11, 7333–7345.

[RPL14] Chiara Risi, Daniel Persson, and Erik G. Larsson, *Massive MIMO with 1-bit ADC*, 2014.

[SDW19] Neev Samuel, Tzvi Diskin, and Ami Wiesel, *Learning to detect*, IEEE Trans. on Sig. Proc. **67** (2019), no. 10, 2554–2564.

[SM21] Mingjie Shao and Wing-Kin Ma, *Binary MIMO detection via homotopy optimization and its deep adaptation*, IEEE Trans. on Sig. Proc. **69** (2021), 781–796.