Estimation, Filtering and Decoding via Deep Learning

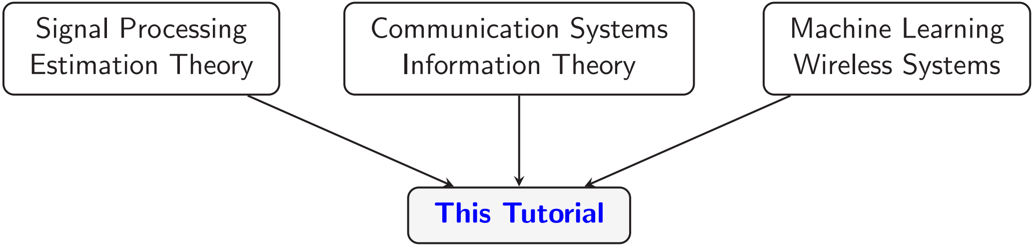
Amir Weiss, Alejandro Lancho and Gary Lee

Massachusetts Institute of Technology, USA

September 4, 2023

Introduction

- Our fields of expertise:



- Connecting the dots: General insights have arisen from different projects
  - US Navy, Task Force Ocean (TFO), Office of Naval Research (ONR)
  - US Department of the Air Force (DAF) MIT AI Accelerator

Acknowledgements

- Collaborators:

Andrew C. Singer

Gregory W. Wornell

Yury Polyanskiy

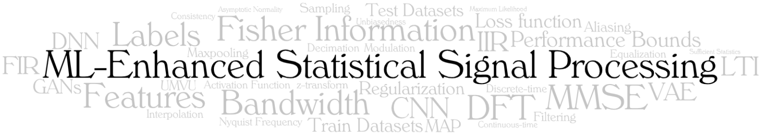
All the TFO project and AIA project team members

- Special thanks to Dr. Binoy Kurien (MIT Lincoln Laboratory Lead)

- Sponsors: MIT, USAF, ONR

Tutorial Goals

- **What is our motivation? Why are we giving this tutorial?**
  - Develop this intermediate, hybrid, but so timely and important emerging field

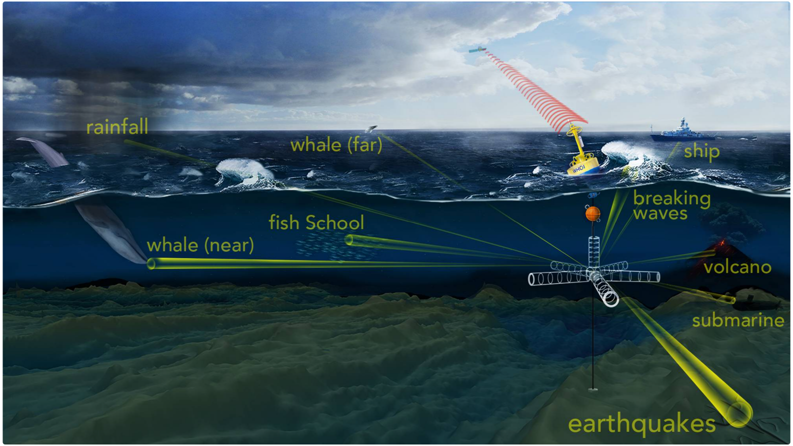


  - Sharing important findings, insights and understanding that are not all published
  - Present and make easily accessible the RF Challenge

- **What do I get from this tutorial?**
  - A methodology for the design of domain-informed DL-based solutions
  - Succinct "rules-of-thumb" for DL-based localization and source separation
  - Access and technical support for the RF Challenge starter code

Motivating Applications: Underwater Acoustic Localization

- Fundamental task in various systems (e.g., harbor defense/monitoring, UUV navigation)

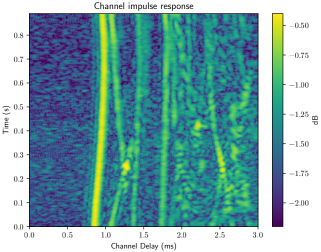


Marine Technology News, March 2021, © Woods Hole Oceanographic Institution, N. Renier

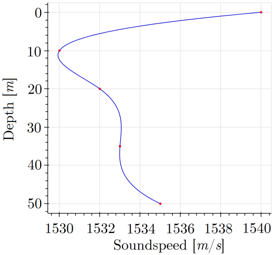
Motivating Applications: Underwater Acoustic Localization

- Fundamental task in various systems (e.g., harbor defense/monitoring, UUV navigation)
- General setting: collection of hydrophones, an acoustic emitter ("source")
- Typical physical characteristics of the underwater acoustic medium:

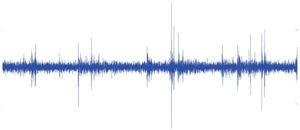
Intricate impulse response



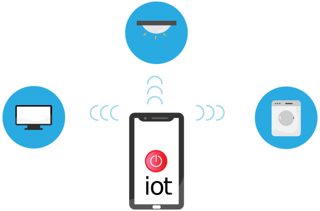
Depth-varying soundspeed profile

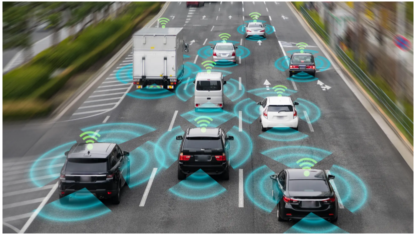


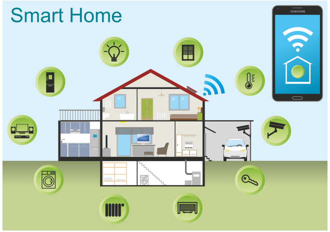
Ambient noise: not Gaussian, not white



- Increasingly congested spectrum $\rightarrow$ more collisions and overlaps
- **Better algorithmic solutions are imperative**



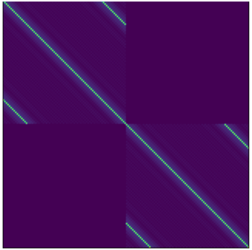




Motivating Applications: RF Signal Separation

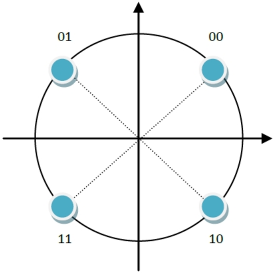
- Increasingly congested spectrum $\rightarrow$ more collisions and overlaps
- **Single-sensor source separation**: key challenge for advanced interference rejection
- **Going beyond stationarity and Gaussianity**, attractive for other problems as well

Nontrivial temporal structures



Temporal covariance matrix
of an OFDM signal

Digital communication:
"discrete" nature



Tutorial Outline

- **Session 1: ML-aided Methodology for Estimation via DNNs**
  - A framework for ML-aided solutions development
  - Underwater Acoustic Localization as a case study

  Speaker: Amir Weiss

- **Session 2: Single-Channel Source Separation of Digital Communication Signals**
  - A communication signal model beyond stationarity
  - DNN source separation performance on digital communication signals

  Speaker: Alejandro Lancho

- **Session 3: Deep Learning Methods, Challenges, and a Short Hands-on Session**
  - On neural architectures for source separation
  - RF Challenge/Hands-on Mini RF Challenge

  Speaker: Gary Lee

Tutorial Outline

- **Session 1: ML-aided Methodology for Estimation via DNNs**
  - A framework for ML-aided solutions development
  - Underwater Acoustic Localization as a case study

  Speaker: Amir Weiss

- **Session 2: Single-Channel Source Separation of Digital Communication Signals**
  - A communication signal model beyond stationarity
  - DNN source separation performance on digital communication signals

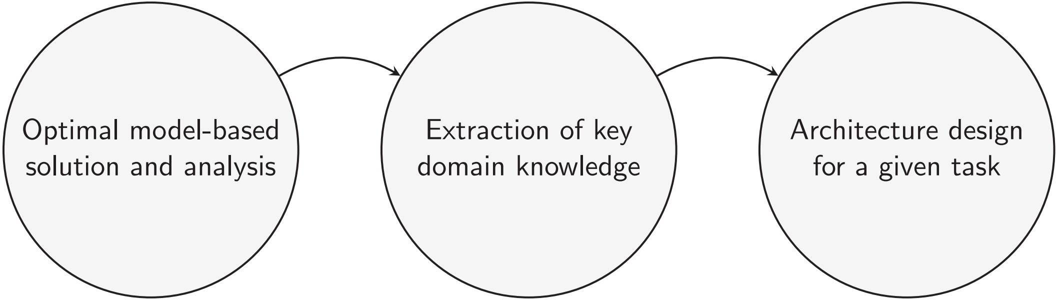
  Speaker: Alejandro Lancho

- **Session 3: Deep Learning Methods, Challenges, and a Short Hands-on Session**
  - On neural architectures for source separation
  - RF Challenge/Hands-on Mini RF Challenge

  Speaker: Gary Lee

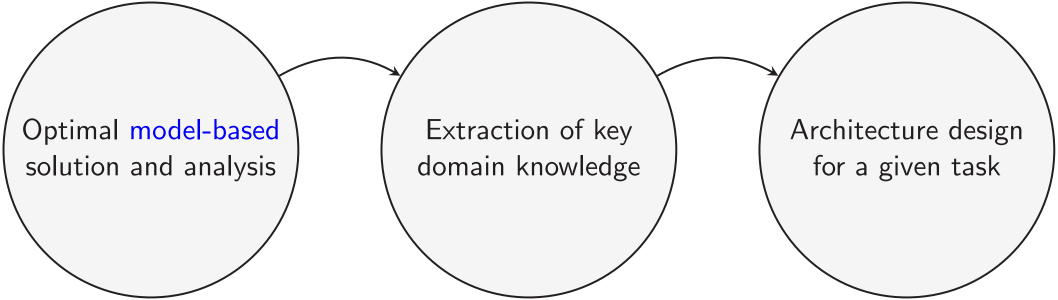
Balancing The "Model-Based" and "Data-Driven" Approaches

- We view the two approaches, rather than contrasting, as complementary
- Design process of an ML-aided solution for a given problem:



Balancing The "Model-Based" and "Data-Driven" Approaches

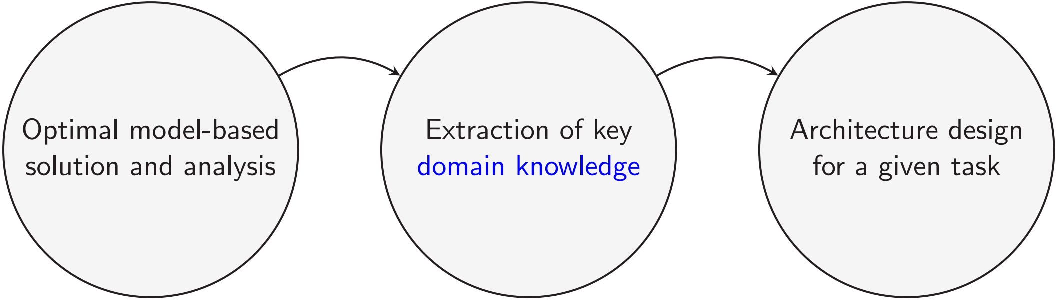
- We view the two approaches, rather than contrasting, as complementary
- Design process of an ML-aided solution for a given problem:



"Compact" representation, amenable for analysis, (can be) easy to interpret

Balancing The "Model-Based" and "Data-Driven" Approaches

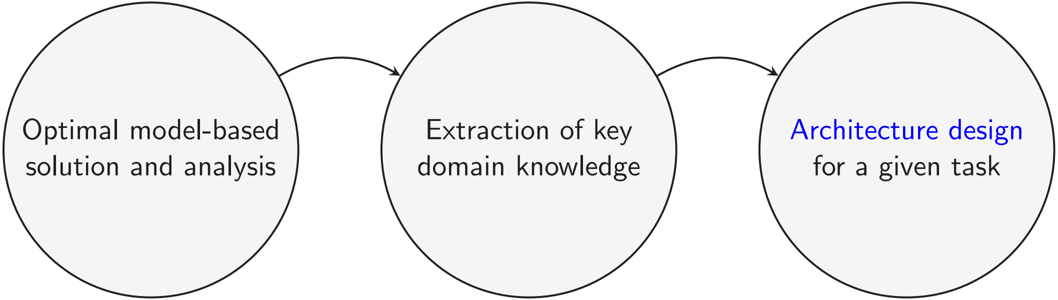
- We view the two approaches, rather than contrasting, as complementary
- Design process of an ML-aided solution for a given problem:



Signals' statistics, physical phenomena, measure of goodness

Balancing The "Model-Based" and "Data-Driven" Approaches

- We view the two approaches, rather than contrasting, as complementary
- Design process of an ML-aided solution for a given problem:

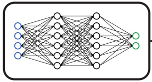


Input structure, key design parameters, training procedure, loss function(s)

Estimation: Localization as a Case Study

- Underwater localization: Enabling technological ability for a variety of applications
- Acoustic waves $\rightarrow$ favorable propagation properties underwater
- Physics of underwater acoustics is (relatively) well-understood
- **However**, analytically complicated $\rightarrow$ classical solutions are typically very limited:
  - High computational load (impractical for online)
  - Require strong prior knowledge about the environment
  - Sensitive to model mismatch

Data-driven methods as a viable solution?

Acoustic data $\longrightarrow$  $\longrightarrow$ Source location

Data-driven direct localization with single-sensor receivers

- Model is generally unknown, **availability of datasets** 
- **Not** a "two-step" method

- **Not** range/azimuth/depth, but an *exact* 3D coordinate
- **Not** arrays (in every receiver), non-coherent processing

Signal Model and Problem Formulation

- Frequency-domain (DFT) baseband of the received signal:

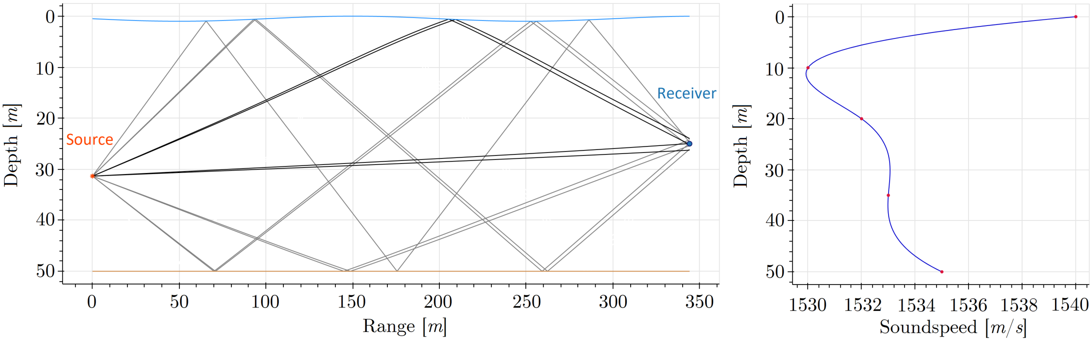
$$\mathbf{x}_\ell = \mathbf{H}_\ell(\mathbf{p}, \mathcal{E})\mathbf{s} + \mathbf{v}_\ell \in \mathbb{C}^{N\times 1}, \qquad \forall \ell \in \{1, \dots, L\}$$

  - $\mathbf{x}_\ell$: received signal at the ℓ-th receiver (observed)
  - $\mathbf{s}$: emitted waveform from the acoustic source (unknown)
  - $\mathbf{v}_\ell$: additive noise, not necessarily Gaussian/white
  - $\mathbf{H}_\ell(\mathbf{p}, \mathcal{E}) = \mathrm{Diag}\left(\mathbf{h}\left(\mathbf{p}, \mathcal{E}\right)\right)$: position- and environment-dependent frequency response
    - $\mathbf{p} \in \mathbb{R}^{3\times 1}$: source's position (unknown $\rightarrow$ our estimand)
    - $\mathcal{E}$: set of environmental parameters (unknown), could be huge

Goal

Given the data $\{\mathbf{x}_\ell\}_{\ell=1}^{L}$, estimate the source's position $\mathbf{p}$

- Illustration of simulated ray propagation model in nonisovelocity environment

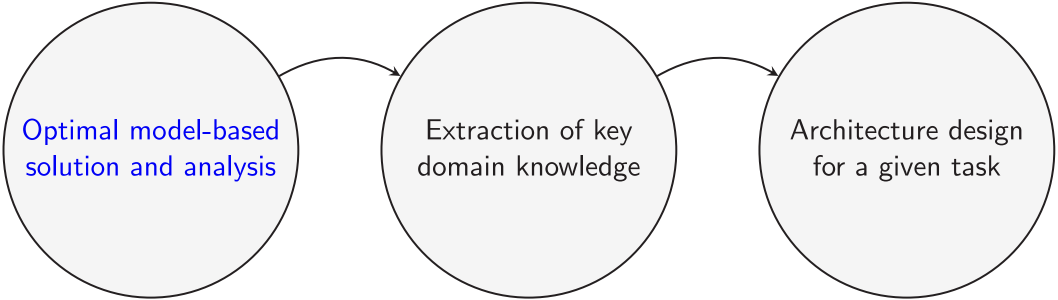


Plots generated using the Bellhop simulator

- Even with realistically simulated data, a **computationally formidable task**

Re-Our Proposed Methodology

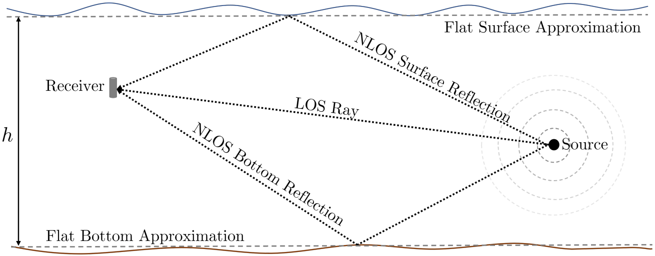
- Design process of an ML-aided solution for a given problem:



"Compact" representation, amenable for analysis, (Can be) Easy to interpret

Step 1: A Model-Based Optimal Solution

- Proposed (tremendously) simplified propagation model: **The 3-ray model**[1]

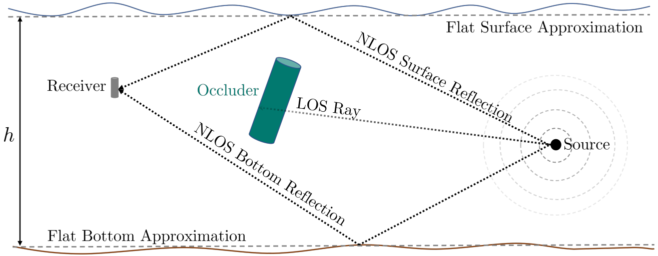


- $$[\mathbf{H}_\ell(\mathbf{p},\mathcal{E})]_{kk} = \underbrace{b_{1\ell}e^{-\jmath\omega_k\tau_{1\ell}(\boldsymbol{p},\mathcal{E})}}_{\text{LOS}} + \underbrace{b_{2\ell}e^{-\jmath\omega_k\tau_{2\ell}(\boldsymbol{p},\mathcal{E})}}_{\text{NLOS, surface}} + \underbrace{b_{3\ell}e^{-\jmath\omega_k\tau_{3\ell}(\boldsymbol{p},\mathcal{E})}}_{\text{NLOS, bottom}}$$

[1] Weiss, A., Arikan, T., Vishnu, H., Deane, G.B., Singer, A.C. and Wornell, G.W., 2022. A semi-blind method for localization of underwater acoustic sources. IEEE Transactions on Signal Processing, 70, pp.3090-3106.

Step 1: A Model-Based Optimal Solution

- Proposed (tremendously) simplified propagation model: **The 3-ray model**[1]



- $$[\mathbf{H}_\ell(\mathbf{p},\mathcal{E})]_{kk} = \underbrace{b_{1\ell}e^{-\jmath\omega_k\tau_{1\ell}(\boldsymbol{p},\mathcal{E})}}_{\text{LOS}} + \underbrace{b_{2\ell}e^{-\jmath\omega_k\tau_{2\ell}(\boldsymbol{p},\mathcal{E})}}_{\text{NLOS, surface}} + \underbrace{b_{3\ell}e^{-\jmath\omega_k\tau_{3\ell}(\boldsymbol{p},\mathcal{E})}}_{\text{NLOS, bottom}}$$
- For example, one special case: $b_{1\ell} = 0$

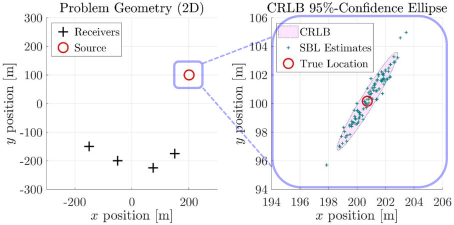
[1] Weiss, A., Arikan, T., Vishnu, H., Deane, G.B., Singer, A.C. and Wornell, G.W., 2022. A semi-blind method for localization of underwater acoustic sources. IEEE Transactions on Signal Processing, 70, pp.3090-3106.

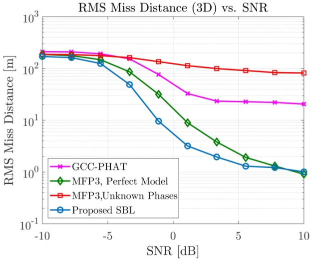
Step 1: A Model-Based Optimal Solution

- Optimal solution[1] (in the least-squares sense):

$$\widehat{\mathbf{p}}_{\mathsf{SBL}} = \underset{\mathbf{p}\in\mathbb{R}^{3\times 1}}{\arg\max}\ \lambda_{\mathsf{max}}\Big(\underbrace{\mathbf{Q}(\mathbf{p},\mathbf{x}_1,\ldots,\mathbf{x}_L)}_{\text{Position- and data-dependent matrix}}\Big)$$

- In white Gaussian noise, for spectrally flat signal, attains the Cramér-Rao lower bound





[1] Weiss, A., Arikan, T., Vishnu, H., Deane, G.B., Singer, A.C. and Wornell, G.W., 2022. A semi-blind method for localization of underwater acoustic sources. IEEE Transactions on Signal Processing, 70, pp.3090-3106.

Step 1: A Model-Based Optimal Solution

- Optimal solution[1] (in the least-squares sense):

$$\widehat{\mathbf{p}}_{\mathsf{SBL}} = \underset{\mathbf{p}\in\mathbb{R}^{3\times1}}{\arg\max}\ \lambda_{\mathsf{max}}\Big(\underbrace{\mathbf{Q}(\mathbf{p},\mathbf{x}_1,\ldots,\mathbf{x}_L)}_{\text{Position- and data-dependent matrix}}\Big)$$

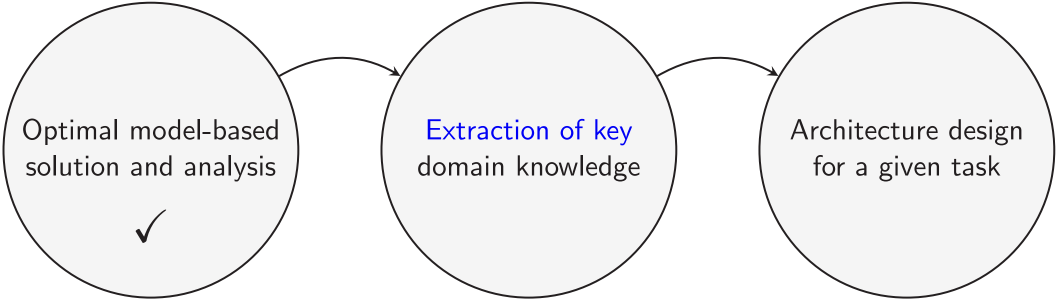
- In white Gaussian noise, for spectrally flat signal, attains the Cramér-Rao lower bound
  - $\longrightarrow$ Trivially extends to an R-ray model with $R > 3$
  - $\longrightarrow$ Extends to an nonisovelocity propagation model
  - $\longrightarrow$ Extends to nonflat ocean surface and bottom

- **What's not good?** Practically, these extensions are **computationally infeasible**
- **What's good?** The method provide a solid **generalizable intuition**!

[1] Weiss, A., Arikan, T., Vishnu, H., Deane, G.B., Singer, A.C. and Wornell, G.W., 2022. A semi-blind method for localization of underwater acoustic sources. IEEE Transactions on Signal Processing, 70, pp.3090-3106.

Re-Our Proposed Methodology

- Design process of an ML-aided solution for a given problem:



Key statistics, physical phenomena, measure of goodness

Step 2: Design Considerations of An ML-Aided Solution

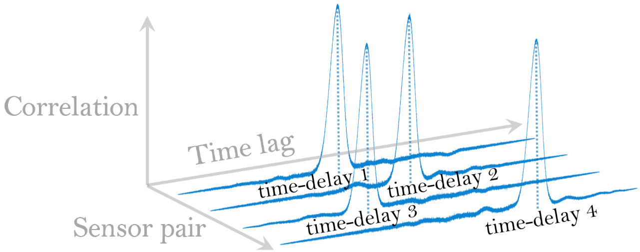
- **Inference computational complexity**: can be reduced?

$$\widehat{\mathbf{p}}_{\mathsf{SBL}} = \underset{\mathbf{p}\in\mathbb{R}^{3\times 1}}{\arg\max}\ \lambda_{\mathsf{max}}\Big(\underbrace{\mathbf{Q}(\mathbf{p},\mathbf{x}_1,\ldots,\mathbf{x}_L)}_{\substack{\text{Position- and data-}\\ \text{dependent matrix}}}\Big)$$

Requires a grid search over a volume of interest + local nonconvex optimization

Step 2: Design Considerations of An ML-Aided Solution

- **Inference computational complexity**: can be reduced?
- If we use a NN for the solution, what **input structure** should it have?
  - How does the above affect the "micro-architectural" choices (e.g., layer type)?



Analysis shows that correlations are key (sufficient statistics)

Step 2: Design Considerations of An ML-Aided Solution

- **Inference computational complexity**: can be reduced?
- If we use a NN for the solution, what **input structure** should it have?
- **Exploit statistical dependencies** between, e.g., azimuth and range?
  - If so, what can be done **in training** to promote such functional behavior?

Non-diagonal Fisher information matrix (e.g., azimuth is informative about range)

Step 2: Design Considerations of An ML-Aided Solution

- **Inference computational complexity**: can be reduced?
- If we use a NN for the solution, what **input structure** should it have?
- **Exploit statistical dependencies** between, e.g., azimuth and range?
- Taking into account the considerations above for this specific domain,

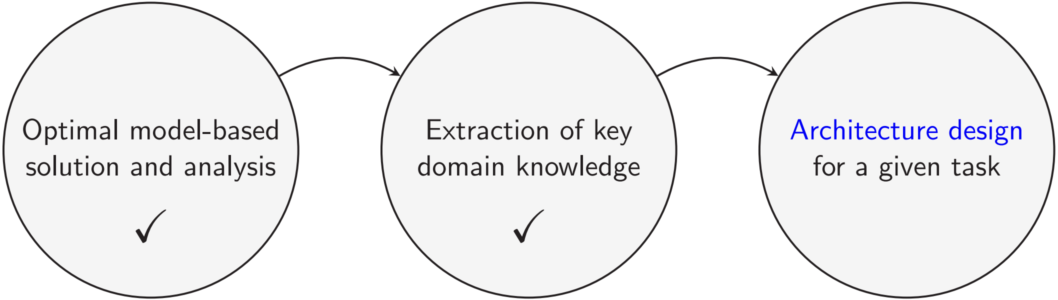
Objective

Obtain a function approximator of an optimal position estimator

*Recall that our **end-goal** is to localize an acoustic source given observed data

Re-Our Proposed Methodology

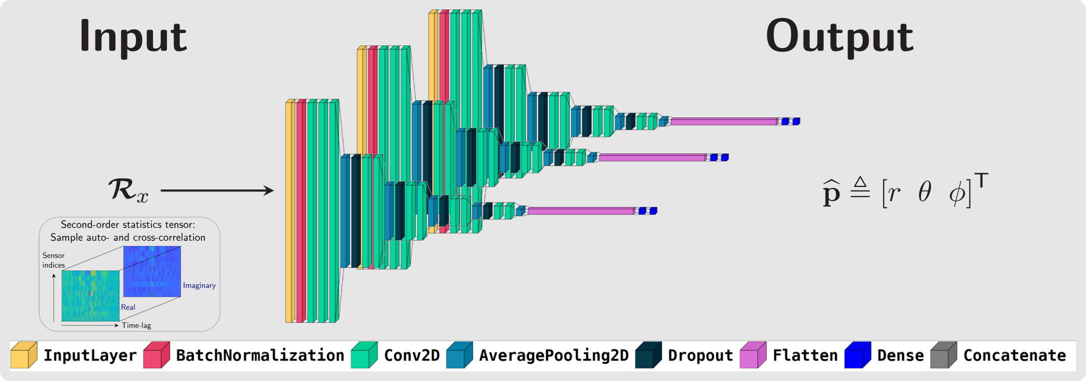
- Design process of an ML-aided solution for a given problem:



Input structure, key design parameters, training procedure, loss function(s)

The Proposed Solution

- **A deep CNN**, input: the SOS tensor, output: position vector in spherical coordinates

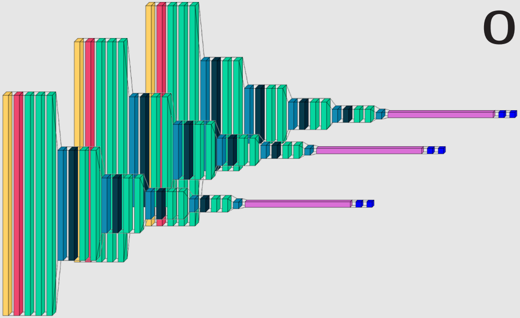


The model is comprised of **three pre-trained sub-models**

- **Inference computational complexity**: can be reduced?

Input **Output**

$\mathcal{R}_x \longrightarrow$

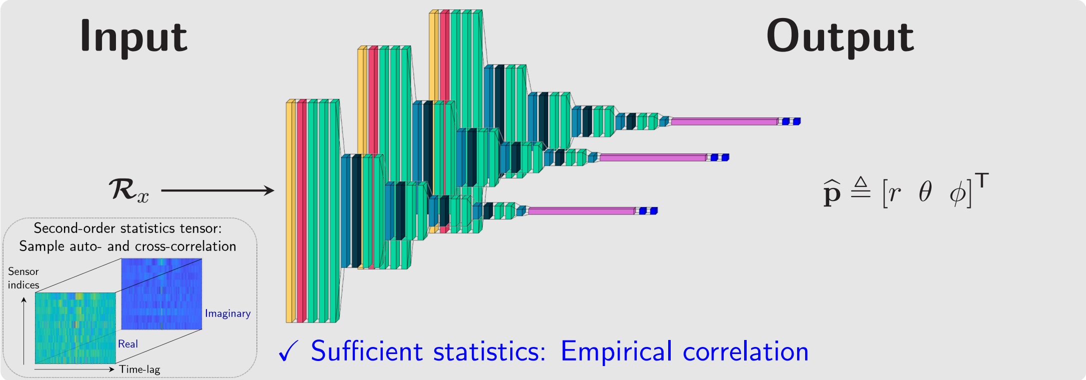


$\widehat{\mathbf{p}} \triangleq [r \quad \theta \quad \phi]^\mathsf{T}$

✓ Overall computational complexity at inference time: fixed

The Proposed Solution

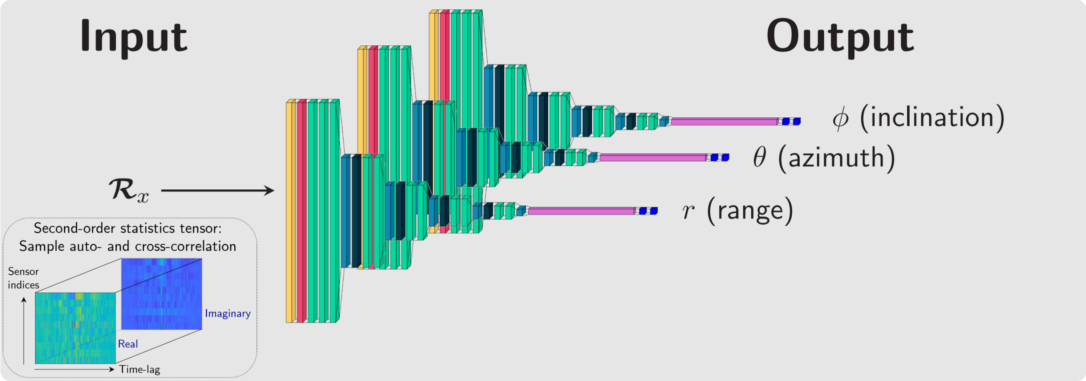
- If we use a NN for the solution, what **input structure** should it have?



✓ Sufficient statistics: Empirical correlation

Conv2D layers + Long kernel size at the first layer

- **Exploit statistical dependencies** between, e.g., azimuth and range?

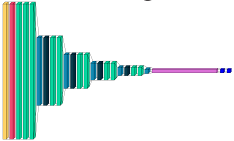


✓ The model is comprised of **three pre-trained sub-models**

Progressive Training and Loss Functions

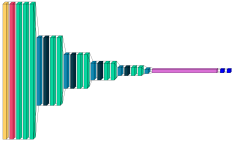
- Phase 1: Train individual models

Range



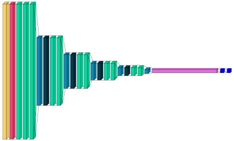
$$\mathcal{L}_r = \left(\widehat{r}(\mathbf{w}_r) - r\right)^2$$

Azimuth



$$\mathcal{L}_\theta \triangleq 2 - 2\cos\left(\widehat{\theta}(\mathbf{w}_\theta) - \theta\right)$$

Inclination



$$\mathcal{L}_\varphi \triangleq 2 - 2\cos\left(2\left(\widehat{\varphi}(\mathbf{w}_\varphi) - \varphi\right)\right)$$

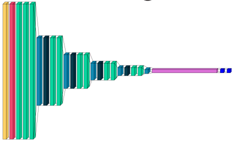
Empirical cyclic error

$$\text{Cyclic Error} = 2 - 2\cos\left(\widehat{\theta} - \theta\right) \overset{|\widehat{\theta}-\theta|<1}{=} \left(\widehat{\theta} - \theta\right)^2 + \mathcal{O}\left(\left(\widehat{\theta} - \theta\right)^4\right) \overset{|\widehat{\theta}-\theta|\ll 1}{\approx} \text{Squared Error}$$

Progressive Training and Loss Functions

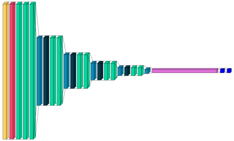
- Phase 1: Train individual models

Range



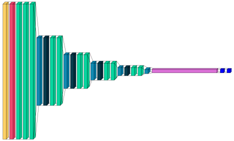
$$\mathcal{L}_r = \left(\widehat{r}(\mathbf{w}_r) - r\right)^2$$

Azimuth



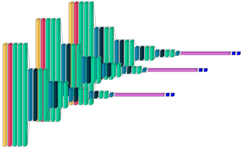
$$\mathcal{L}_\theta \triangleq 2 - 2\cos\left(\widehat{\theta}(\mathbf{w}_\theta) - \theta\right)$$

Inclination



$$\mathcal{L}_\varphi \triangleq 2 - 2\cos\left(2\left(\widehat{\varphi}(\mathbf{w}_\varphi) - \varphi\right)\right)$$

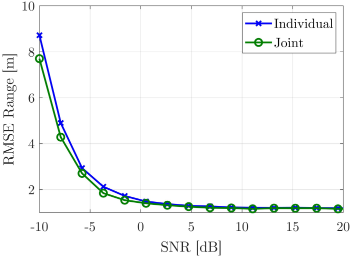
- Phase 2: Train global model with “hot” initialization (joint optimization approximator)



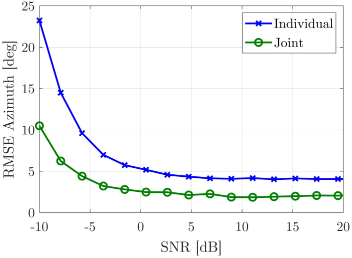
$$\|\widehat{\mathbf{p}}(\mathbf{w}_p) - \mathbf{p}\|_2^2 = r^2 + \widehat{r}^2(\mathbf{w}_p) - 2r\widehat{r}(\mathbf{w}_p)\left[\sin(\theta)\sin\left(\widehat{\theta}(\mathbf{w}_p)\right)\cos\left(\varphi - \widehat{\varphi}(\mathbf{w}_p)\right) + \cos(\theta)\cos\left(\widehat{\theta}(\mathbf{w}_p)\right)\right]$$

- 3-ray propagation, individual DNN models vs. global DNN model
- $L = 4$ sensors, $N = 100$ samples, (constant) speed of sound $c = 1500\frac{m}{s}$, depth $50m$

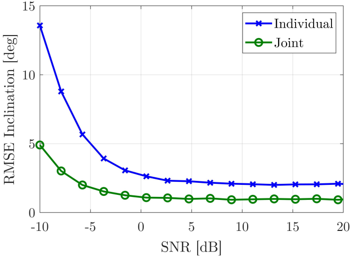
Range



Azimuth

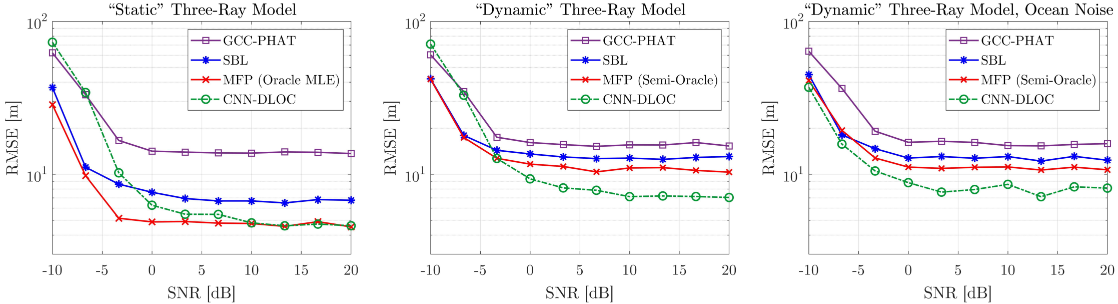


Inclination



- As expected, accuracy (uniformly) higher in joint estimation (/direct localization)

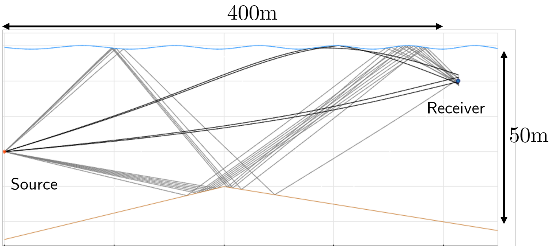
Simulation Results: 3-Ray Model

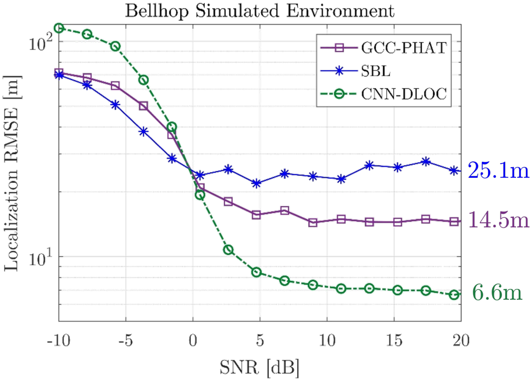


Matching the oracle optimal solution
that has access to prior environmental knowledge

Random perturbations added to the "surface-ray" time-delays
More robust relative to model-based alternatives

Simulation Results: Bellhop Simulated Environment

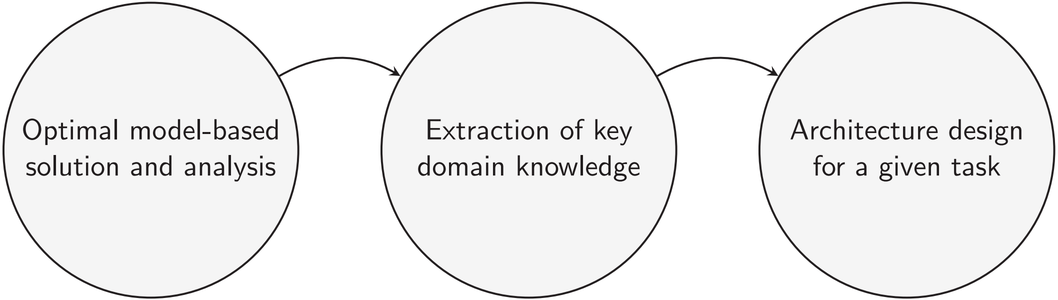




- **Depth-varying speed of sound**
- Undulating surface, **small seamount** at bottom
- **Superior performance, faster inference computation time**

Estimation via DNNs: General Takeaways

- Systematic development approach to ML-aided data-driven solutions



- Key architectural choices—naturally arises from classical signal processing (SP) concepts:
  - Input structure (via the notion of sufficient statistics)
  - NN architecture (informed by basic SP operations, such as filtering)
  - Loss functions (some are well-known in SP literature, e.g., cyclic error)
  - Training procedure (analogy to iterative algorithms)

Tutorial Outline

- **Session 1: ML-aided Methodology for Estimation via DNNs**
  - A framework for ML-aided solutions development
  - Underwater Acoustic Localization as a case study

  Speaker: Amir Weiss

- **Session 2: Single-Channel Source Separation of Digital Communication Signals**
  - A communication signal model beyond stationarity
  - DNN source separation performance on digital communication signals

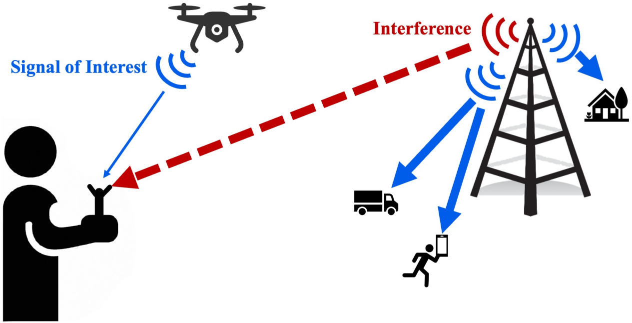
  Speaker: Alejandro Lancho

- **Session 3: Deep Learning Methods, Challenges, and a Short Hands-on Session**
  - On neural architectures for source separation
  - RF Challenge/Hands-on Mini RF Challenge

  Speaker: Gary Lee

Motivation

- Radio spectrum increasingly crowded $\longrightarrow$ **spectrum sharing unavoidable**
  - To keep **high reliabilities** $\longrightarrow$ **signal separation** essential module



- Gives rise to a **source separation problem**

Problem Setup I

- We consider single-antenna receivers $\Rightarrow$ no spatial diversity

  $\Rightarrow$ Single-channel source separation (**SCSS**)

$$\mathsf{y}[n] = \mathsf{s}[n - \mathsf{k}_s] + \rho_{\mathsf{SIR}}^{-1/2}\mathsf{b}[n - \mathsf{k}_b] + \rho_{\mathsf{SNR}}^{-1/2}\mathsf{w}[n], \ \ n \in \mathbb{Z}$$

  - $\mathsf{s}[n]$, $\mathsf{b}[n]$: signal of interest (SOI) and interference, resp.

    $\rightarrow$ statistically independent

  - $\rho_{\mathsf{SIR}}, \ \rho_{\mathsf{SNR}} \in \mathbb{R}_+$, (SIR: Signal-to-interference ratio)

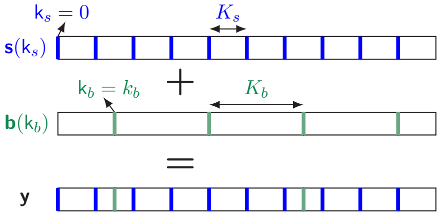
- For a recording of N samples:

$$\mathbf{y} = \mathbf{s}(\mathsf{k}_s) + \rho_{\mathsf{SIR}}^{-1/2}\mathbf{b}(\mathsf{k}_b) + \rho_{\mathsf{SNR}}^{-1/2}\mathbf{w}$$

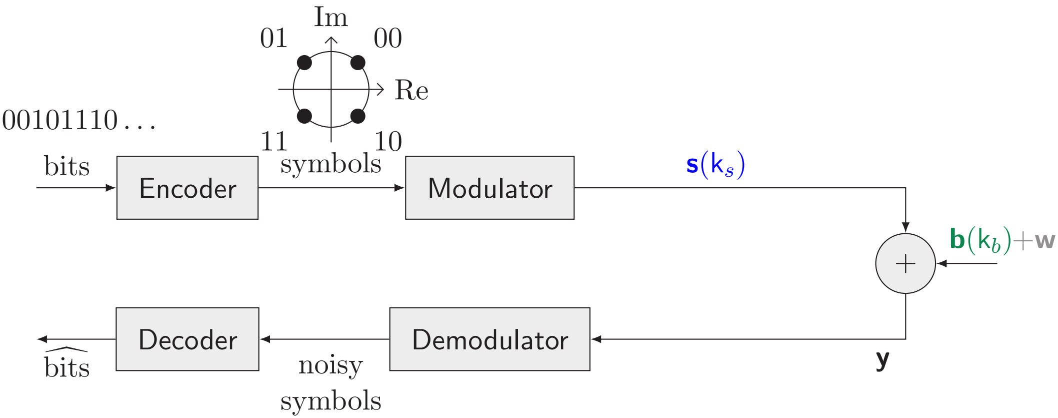
$$\mathbf{y} = \mathbf{s}(\mathsf{k}_s) + \rho_{\mathsf{SIR}}^{-1/2}\mathbf{b}(\mathsf{k}_b) + \rho_{\mathsf{SNR}}^{-1/2}\mathbf{w}$$

- $\mathbf{s}(\mathsf{k}_s), \mathbf{b}(\mathsf{k}_b)$ assumed to be zero-mean, unit-variance, **cyclostationary signals**
  - K_s, K_b: cyclic periods
  - k_s, k_b: **arbitrary time shifts** w.r.t. start of cyclic period of $\mathsf{s}[n]$, $\mathsf{b}[n]$

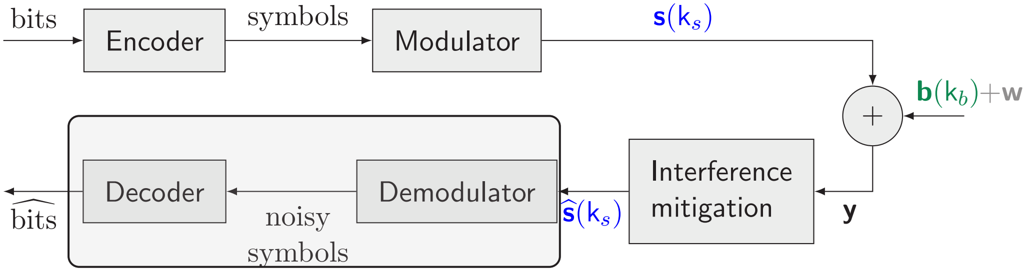
    $\rightarrow \mathsf{k}_s \sim \mathrm{Unif}\{1, \ldots, K_s\}$, $\mathsf{k}_b \sim \mathrm{Unif}\{1, \ldots, K_b\}$



Scheme of a Digital Communication System

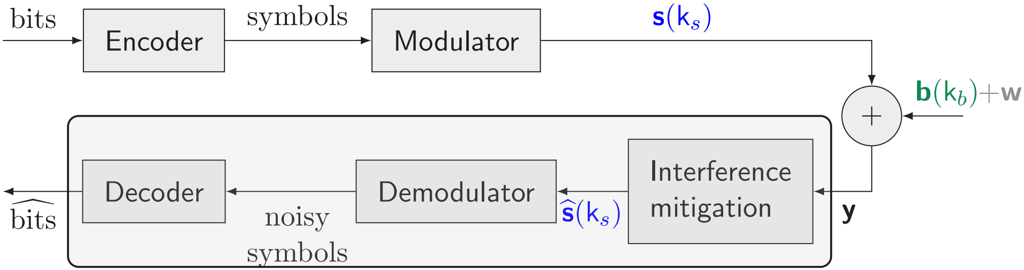


Scheme of a Digital Communication System



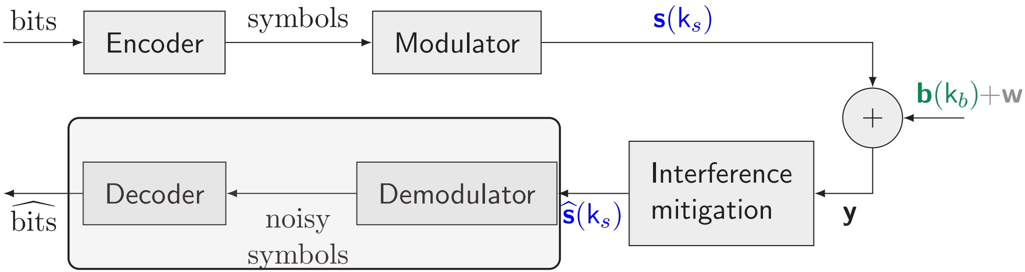
separation + demodulation

Scheme of a Digital Communication System



end-to-end demodulation

Scheme of a Digital Communication System



- **This session**: Separation + demodulation
- Figures of merit: **MSE** and bit-error rate (**BER**) as a function of **SIR**

Traditional Interference Rejection in Communication Systems

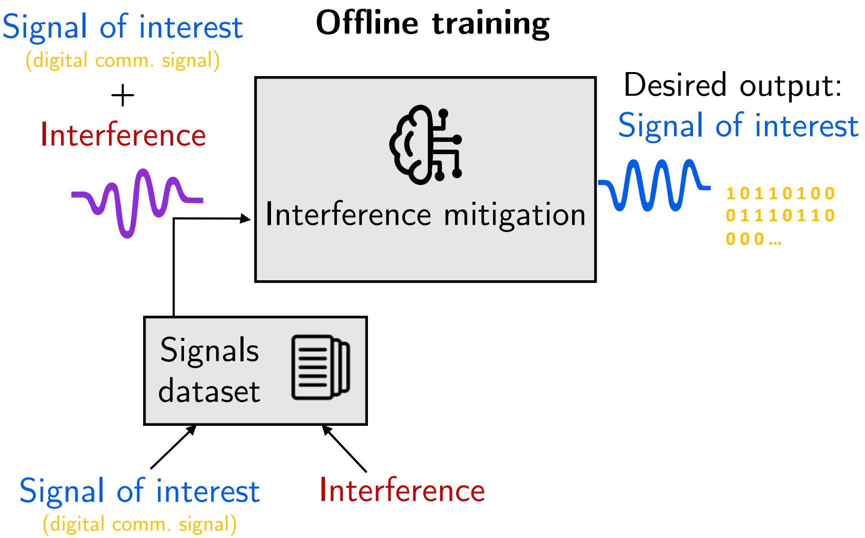
- **Signal detection**:
  - **Matched filtering**: Optimal in detection (SNR sense) when interference is Gaussian
- **Signal estimation**:
  - Optimal linear estimator in MSE sense (**LMMSE**) (not necessarily the MMSE):

$$\widehat{\mathbf{s}} = \mathbf{C}_{ss}(\mathbf{C}_{ss} + \mathbf{C}_{vv})^{-1}\,\mathbf{y}$$

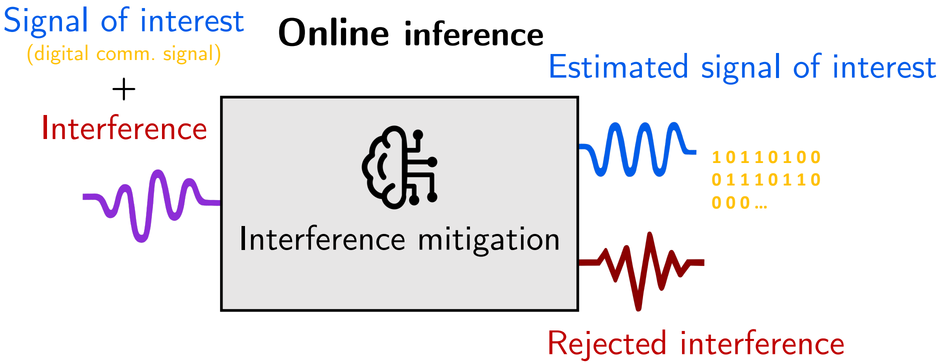
$\mathbf{C}_{ss}, \mathbf{C}_{vv}$: Covariance matrices of $\mathbf{s}(\mathsf{k}_s)$ and $\mathbf{v}(\mathsf{k}_b) \triangleq \rho_{\mathsf{SIR}}^{-1/2}\mathbf{b}(\mathsf{k}_b) + \rho_{\mathsf{SNR}}^{-1/2}\mathbf{w}$, resp.

- **Potential problem**: They can be applied in different ("small") time scales
  - **The longer the better**
  - **The longer the more complex**

A Data-Driven Pipeline for Interference Mitigation: Training



A Data-Driven Pipeline for Interference Mitigation: Inference



Motivation to Pursue a Data-Driven Approach

- Deep neural networks (DNNs) successful for source separation
  - **Computer vision**: Color features and local features
  - **Audio**: Spectogram masking methods

Many communication signals

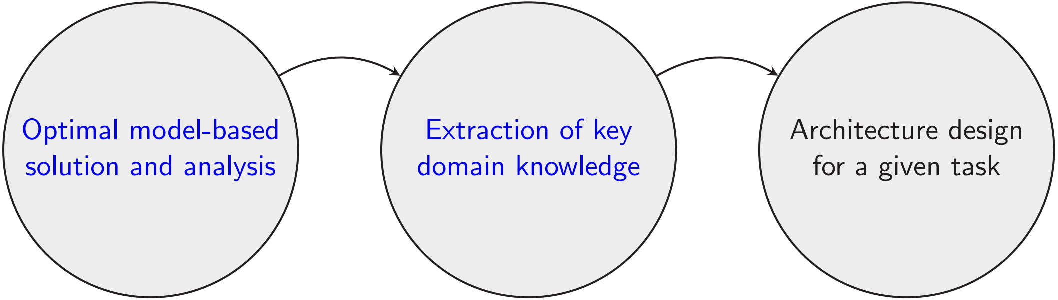
- May not have local features/dependencies (e.g. OFDM signals)
- Overlapping in time and frequency

$\Rightarrow$ **Domain-specific knowledge is needed for successful operation**

From Model-Based to Data-Driven

- When **prior knowledge** on the signal models is **not known or available**
  - **Model-based solution becomes infeasible**
  - **We can still learn from simplified signal models**

    $\longrightarrow$ From model-based to data-driven solutions:



A Signal Model Beyond Stationarity

- **Cyclostationary Gaussian mixture** model:

$$\mathbf{y} = \mathbf{s}(\mathsf{k}_s) + \rho_{\mathsf{SIR}}^{-1/2}\mathbf{b}(\mathsf{k}_b) + \rho_{\mathsf{SNR}}^{-1/2}\mathbf{w} = \mathbf{s}(\mathsf{k}_s) + \mathbf{v}(\mathsf{k}_b)$$

  - $\mathbf{s}(m_s) \sim \mathcal{CN}(\mathbf{0}, \mathbf{C}_{ss}(m_s))$, $\mathbf{b}(m_b) \sim \mathcal{CN}(\mathbf{0}, \mathbf{C}_{bb}(m_b))$, $\mathbf{w} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$
  - K_s, K_b: cyclic periods
  - k_s, k_b: **arbitrary time shifts** w.r.t. start of cyclic period of $\mathsf{s}[n]$, $\mathsf{b}[n]$

    $\rightarrow \mathsf{k}_s \sim \mathrm{Unif}\{1, \ldots, K_s\}$, $\mathsf{k}_b \sim \mathrm{Unif}\{1, \ldots, K_b\}$
  - ρ_{SIR}, $\rho_{\mathsf{SNR}} \in \mathbb{R}_+ \rightarrow$ **assumed to be known/ can be estimated**

Objective

Obtain understanding based on analysis

$\Rightarrow$ **Make informed architectural decisions**

Assume Models of $\mathbf{s}(\mathsf{k}_s)$ and $\mathbf{b}(\mathsf{k}_b)$ Are Known

- Signals $\mathbf{y}$ and $\mathbf{s}$ jointly Gaussian $\Rightarrow$ **optimal estimator can be easily derived**:

$$\widehat{\mathbf{s}}_{\mathsf{MMSE}} = \mathbb{E}\left[\mathbb{E}[\mathbf{s}(\mathsf{k}_s)|\mathbf{y}, \mathsf{k}_s, \mathsf{k}_b]|\mathbf{y}\right] = \mathbb{E}\left[\widehat{\mathbf{s}}_{\mathsf{CLMMSE}}(\mathsf{k}_s, \mathsf{k}_b)|\mathbf{y}\right]$$
$$= \sum_{m_s=1}^{K_s} \sum_{m_b=1}^{K_b} \mathbb{P}[\mathsf{k}_s = m_s, \mathsf{k}_b = m_b|\mathbf{y}]\, \widehat{\mathbf{s}}_{\mathsf{CLMMSE}}(m_s, m_b)$$

with

$$\widehat{\mathbf{s}}_{\mathsf{CLMMSE}}(m_s, m_b) \triangleq \mathbf{C}_{sy}(m_s, m_b)\mathbf{C}_{yy}(m_s, m_b)^{-1}\mathbf{y}$$
$$= \mathbf{C}_{ss}(m_s)\left(\mathbf{C}_{ss}(m_s) + \mathbf{C}_{vv}(m_b)\right)^{-1}\mathbf{y}$$

- Although $\widehat{\mathbf{s}}_{\mathsf{MMSE}}$ is **computable**, it becomes **harder as signal length grows**

Is the Previous Assumption Reasonable?

- **Three main problems**:
  - **Computing** $\mathbb{P}[\mathsf{k}_s = m_s, \mathsf{k}_b = m_b | \mathbf{y}]\ \forall m_s, m_b$ computationally **involved** as $K_s, K_b \uparrow$
  - $(\mathbf{C}_{ss} + \mathbf{C}_{vv})^{-1}$ involves a **large matrix inversion** (for long observations)
  - If signal model is not given $\Rightarrow$ covariance matrices are not given

    $\Rightarrow$ **Estimating covariance matrix requires dataset synchronization**

An Approach if Synchronized Datasets: MAP-QLMMSE

- **Assumption**: Covariance matrices can be estimated from syncrhonized dataset
- **Two step synchronization-separation**:
  - **MAP estimation of time shift**: $\widehat{\mathsf{k}}_b^{\mathsf{MAP}} \triangleq \arg\max_{m\in\{1,\dots,K_b\}} \mathbb{P}[\mathsf{k}_b = m|\mathbf{y}]$
  - **MAP-based quasi-LMMSE estimator**: $\widehat{\mathbf{s}}_{\mathsf{MAP\text{-}QLMMSE}} \triangleq \widehat{\mathbf{s}}_{\mathsf{LMMSE}}\big(\widehat{\mathsf{k}}_b^{\mathsf{MAP}}\big)$
- We show that the MAP-QLMMSE estimator is **asymptotically optimal**[2]
  - Under mild condition (shift uniquely detectable):

$$\mathbb{P}\left[\widehat{\mathsf{k}}_b^{\mathsf{MAP}} \neq \mathsf{k}_b\right] = o\left(\frac{1}{N^\alpha}\right), \qquad \lim_{N\to\infty} \frac{\mathbb{E}\left[\|\widehat{\mathbf{s}}_{\mathsf{MMSE}} - \mathbf{s}\|_2^2\right]}{\mathbb{E}\left[\|\widehat{\mathbf{s}}_{\mathsf{MAP\text{-}QLMMSE}} - \mathbf{s}\|_2^2\right]} = 1$$

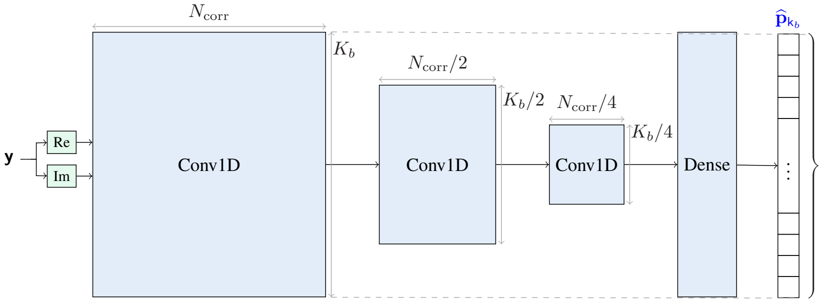
- **MAP estimator computationally hard**

[2] A. Lancho, A. Weiss, G. Lee, J. Tang, Y. Bu, Y. Polyanskiy, and G. Wornell, "Data-Driven Blind Synchronization and Interference Rejection for Digital Communication Signals," IEEE GLOBECOM, Rio de Janeiro, Brazil, 2022.

MAP-QLMMSE Implementation: CNN-QLMMSE

- Assume **availability of dataset** $\left\{\mathbf{s}(\mathsf{k}_s = 0)^{(i)}, \mathbf{b}(\mathsf{k}_b = k_b)^{(i)}\right\}$, $i \in \{1, \ldots, T\}$

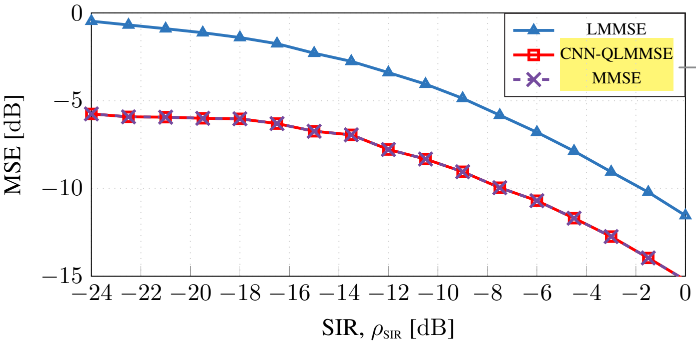
  $\Rightarrow$ **Data-driven approach** to estimate $\widehat{\mathsf{k}}_b^{\mathsf{MAP}}$ via CNNs $\longrightarrow$ **CNN-QLMMSE**



estimated vector of posterior probabilities

- LMMSE, MMSE and CNN-QLMMSE estimators for fixed SNR (ρ_{SNR}) of $20\,\text{dB}$

$$\mathbf{y} = \underbrace{\mathbf{s}(\mathsf{k}_s = 0)}_{\mathbf{s}(\mathsf{k}_s=0)\sim\mathcal{CN}(\mathbf{0},\mathbf{C}_{ss}(0))} + \underbrace{\rho_{\text{SIR}}^{-1/2}\mathbf{b}(\mathsf{k}_b) + \rho_{\text{SNR}}^{-1/2}\mathbf{w}}_{\substack{\mathbf{v}(\mathsf{k}_b) \\ \mathbf{v}(\mathsf{k}_b)\big|\mathsf{k}_b\sim\mathcal{CN}(\mathbf{0},\mathbf{C}_{vv}(\mathsf{k}_b)), \\ \mathsf{k}_b\sim\text{Unif}\{1,\dots,K_b\}}} \in \mathbb{C}^{320\times 1}$$



→ **Both coincide**

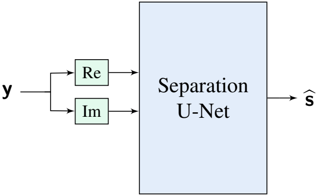
Revisiting the Assumptions Made

- **Three main problems**:
  - **Computing** $\mathbb{P}[\mathsf{k}_s = m_s, \mathsf{k}_b = m_b | \mathbf{y}] \ \forall m_s, m_b$ computationally **involved** as $K_s, K_b \uparrow$
  - $(\mathbf{C}_{ss} + \mathbf{C}_{vv})^{-1}$ involves a **large matrix inversion** (for long observations)
  - If signal model not given $\Rightarrow$ covariance matrices not given

    $\Rightarrow$ **Estimating covariance matrix requires dataset synchronization**

Can a DNN be Competitive and Overcome All Difficulties?

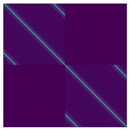
- **Assumption:** Synchronized dataset not available



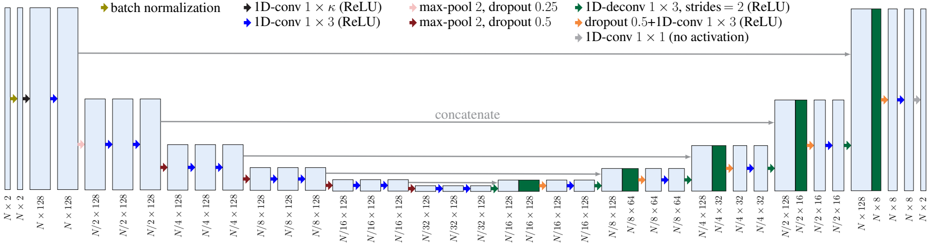
Main domain knowledge modification

Long kernel first layer

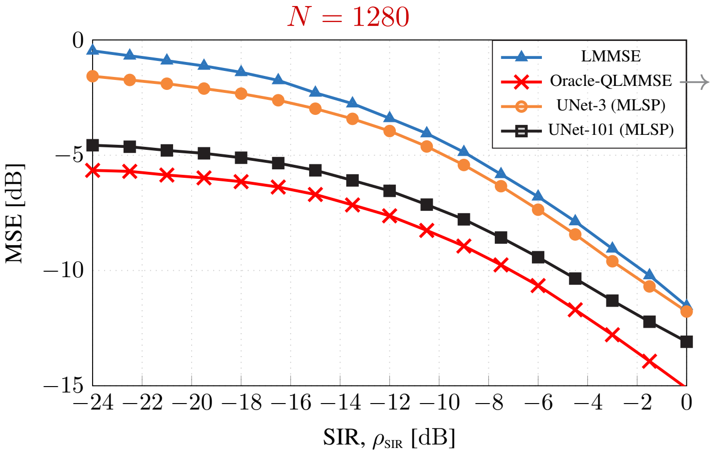
Temporal covariance of OFDM signal:



- U-Net architecture:



Simulation Results on Short Gaussian Mixture[3]



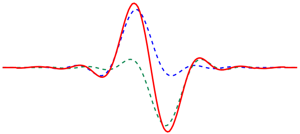
→ Requires oracle access to synchronized statistics and correct shift

[3] G.C.F. Lee, A. Weiss, A. Lancho, J. Tang, Y. Bu, Y. Polyanskiy and G.W. Wornell, "Exploiting temporal structures of cyclostationary signals for data-driven single-channel source separation," in Proc. IEEE International Workshop for Machine Learning and Signal Processing (MLSP), Xi'an, China, Aug. 2022. (**Best student paper award**)

Beyond Gaussianity: Digital Communication Signals

$$\mathbf{y} = \mathbf{s}(\mathsf{k}_s = 0) + \rho_{\mathsf{SIR}}^{-1/2}\mathbf{b}(\mathsf{k}_b) + \rho_{\mathsf{SNR}}^{-1/2}\mathbf{w}$$

- $\mathbf{s}(\mathsf{k}_s = 0)$ bears **QPSK symbols** using a **root-raised cosine pulse-shaping filter**
  - Spanning 8 symbols, oversampling factor $= 16$

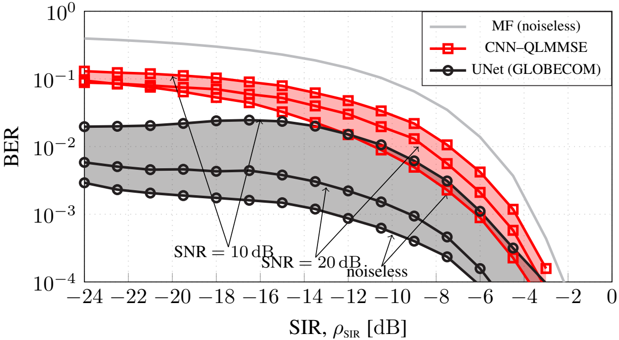


- $\mathbf{b}(\mathsf{k}_b)$ bears **16QAM OFDM symbols** of length $K_b = 80$ with $\mathsf{k}_b \sim \mathrm{Unif}\{1, \ldots, K_b\}$
  - FFT size $= 64$, cyclic-prefix length $= 16$
- Details on signals generation process $\longrightarrow$ **Visit our Github repository**:

  https://github.com/RFChallenge/SCSS_Sync
- **Performance metric**: BER
  - Every approach includes **a last (standard) MF step prior to hard decoding**

Simulation Results for Different SNR Values[4]

- Input length:
  - Separation UNet: $N = 40960$
  - CNN-QLMMSE: $N = 320$



[4] A. Lancho, A. Weiss, G. Lee, J. Tang, Y. Bu, Y. Polyanskiy, and G. Wornell, "Data-Driven Blind Synchronization and Interference Rejection for Digital Communication Signals," IEEE GLOBECOM, Rio de Janeiro, Brazil, 2022.

Summary Session 2

- Learned that from the **cyclostationary Gaussian mixture model**
  - **Synchronization** (dataset level or shifts) $\longrightarrow$ significant **performance gains**
  - **End-to-end DNN** architecture needs to be able to **synchronize** (even implicitly)
    $\longrightarrow$ **UNet** + long kernel first layer (important for capturing temporal structures)
- **Why UNet is a good architectural solution?**
  - What are the **specific characteristics** suitable to our problem?
  - Are these factors **necessary or sufficient**?
  - Would **other modern DNN architectures** work as well?

    $\longrightarrow$ **Answers in the next session**

Instructions for Session 3/Hands-On Session

- Link to the Mini RF Challenge:
  `https://www.kaggle.com/competitions/mini-rf-challenge`

  → Go to Code > Hands-On Session

  or

- Link to Hands-On Session Notebook:
  `https://www.kaggle.com/code/garycflee/hands-on-session/notebook`

- Ensure you are signed into Kaggle.
- In the Hands-On Session notebook, click **"Copy and Edit"**.

- *(For those with issues on Kaggle, you can try the Google Colab Link:*
  `https://bit.ly/RFHandsOn2023`*)*

Tutorial Outline

- **Session 1: ML-aided Methodology for Estimation via DNNs**
  - A framework for ML-aided solutions development
  - Underwater Acoustic Localization as a case study

  Speaker: Amir Weiss

- **Session 2: Single-Channel Source Separation of Digital Communication Signals**
  - A communication signal model beyond stationarity
  - DNN source separation performance on digital communication signals

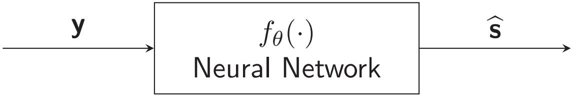
  Speaker: Alejandro Lancho

- **Session 3: Deep Learning Methods, Challenges, and a Short Hands-on Session**
  - On neural architectures for source separation
  - RF Challenge/Hands-on Mini RF Challenge

  Speaker: Gary Lee

SCSS as a Multivariate Regression Problem

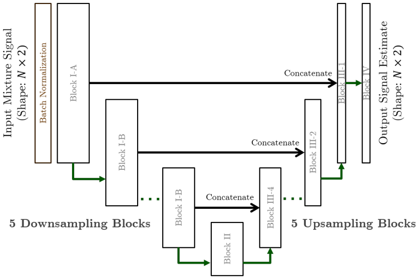
- Learning an end-to-end separator
- Appropriate Parameterization $\Rightarrow$ Neural Architecture Choice



$$\arg\min_{\theta} \mathbb{E}_{\mathbf{y},\mathbf{s}} \left\{ \| f_\theta(\mathbf{y}) - \mathbf{s} \|_2^2 \right\}$$

Minimum Mean-Square Error Estimator

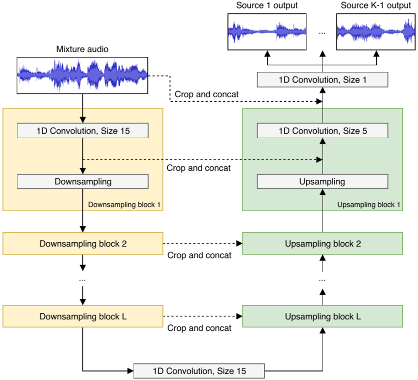
U-Net Architecture



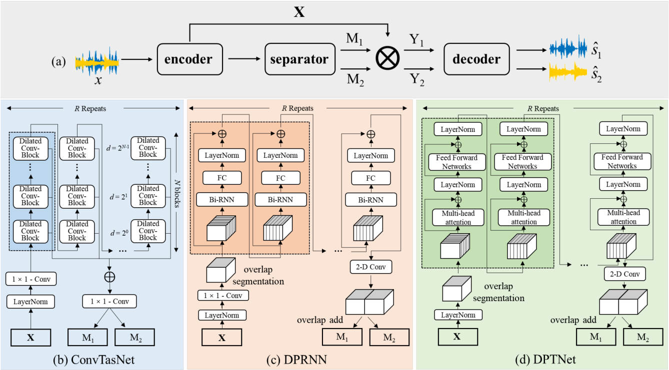
- A fully convolutional network architecture with the same input and output size
- First used in biomedical image segmentation[5]
- Successive downsampling and upsampling blocks (multiresolution features)

[5] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," MICCAI 2015. Lecture Notes in Computer Science. Springer International Publishing, pp. 234–241, 2015.

Other Neural Architectures (from Audio Separation)



Wave-U-Net[6]



TasNet[7]

---

6 Figure from D. Stoller, S. Ewert, and S. Dixon, "Wave-U-Net: A Multi-Scale Neural Network for End-to-End Audio Source Separation," arXiv:1806.03185 [cs.SD], Jun. 2018.

7 Figure from H. Li, K. Chen, L. Wang, J. Liu, B. Wan, and B. Zhou, "Sound source separation mechanisms of different deep networks explained from the perspective of auditory perception," Applied Sciences, vol. 12, no. 2, 2022

Comparing Neural Architectures—A New Baseline

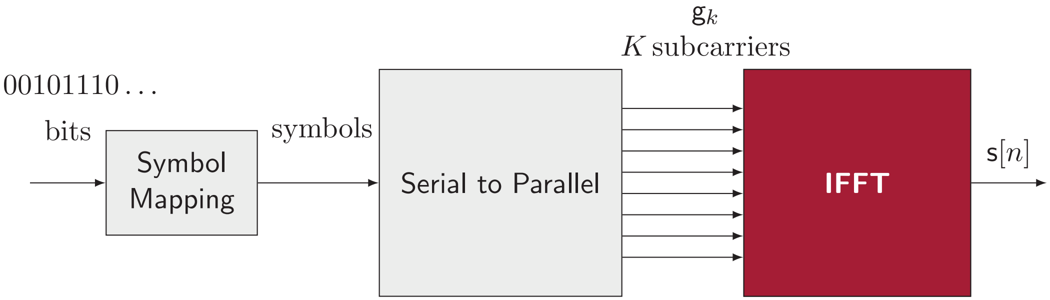
- Which neural architecture should I use? Does it matter?
  - Appropriate architecture leads to more efficient training/better generalization
  - Architectures are typically chosen based on precedence, intuition, and trial and error
  - *What works for image/audio might not work for RF signals (?)*
- Demonstration with Separating OFDM Structures *(Simple Problem Abstraction)*

"Special Case": Separating Multiple-Access-like OFDM Symbols

- Real-valued time-domain signals, representative of **RF signals**
- **Perfect separation** is theoretically attainable *if* source models were known
- Unable to separate with second-order structures alone
- OFDM/Fourier parameters are not explicitly provided
  (i.e., have to be ***learned from data****)*

Comparing Neural Architectures—Separating OFDM

- OFDM Generative Pipeline



$$\mathsf{s}[n] = \sum_{k=0}^{K-1} \mathsf{g}_k \exp\left(j2\pi kn/K\right)$$

Comparing Neural Architectures—Separating OFDM

(Unobserved)

$$\mathsf{s}[n] = \sum_{k=0}^{K-1} \mathsf{g}_k \, r[n - T_{\mathrm{cp}}, \, k] \; , \; \mathsf{b}[n] = \sum_{k=0}^{K-1} \mathsf{h}_k \, r[n - T_{\mathrm{cp}}, \, k],$$

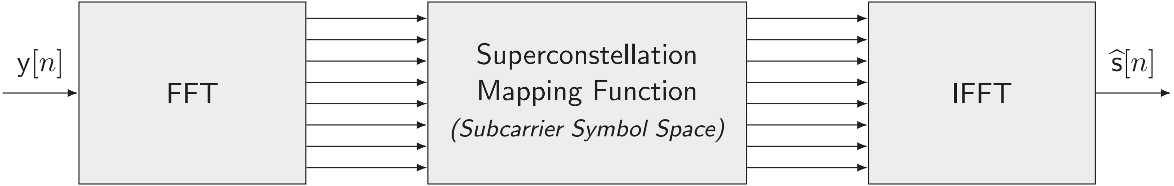
$$r[n, \, k] \triangleq \exp\left(j 2\pi k n / K\right) \mathbb{1}_{\{-T_{\mathrm{cp}} \leq n < K\}},$$

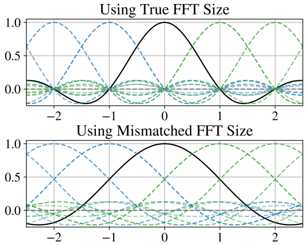
$$\mathsf{y}[n] = \mathsf{s}[n] + \mathsf{b}[n] = \sum_{k=0}^{K-1} a_k \, \exp\left(j 2\pi k (n - T_{\mathrm{cp}}) / K\right) \mathbb{1}_{\{0 \leq n < K + T_{\mathrm{cp}}\}},$$

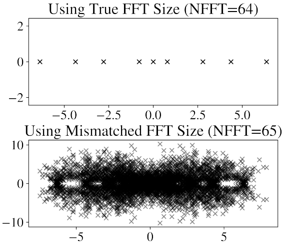
$$\mathsf{a}_k = \mathsf{g}_k + \mathsf{h}_k \; , \; \mathsf{a}_k \in \mathcal{A}.$$

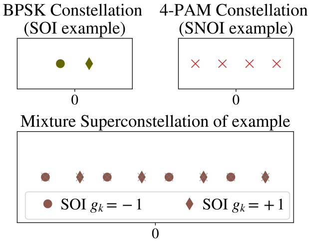
Goal: Estimate s from observation y.

Separating OFDM: (Oracle) Model-Based Approach

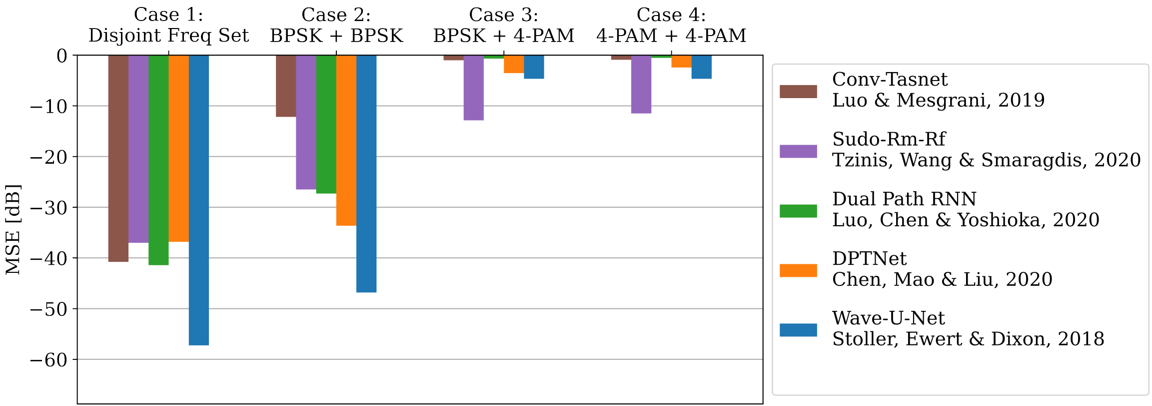






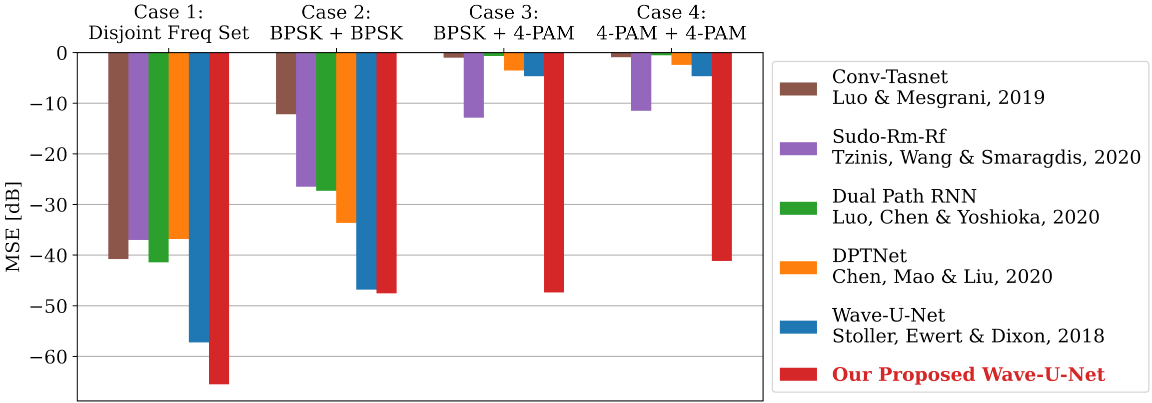


Comparing Neural Architectures—Separating OFDM



Audio-Oriented NN Architectures perform poorly in Cases 3 and 4.

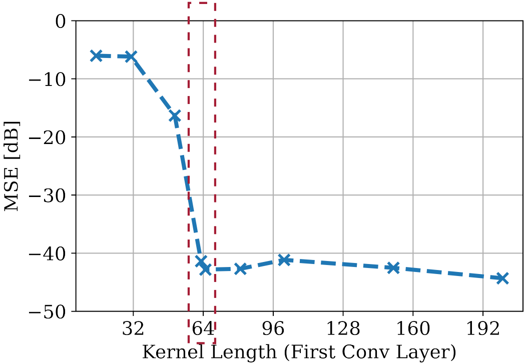
Comparing Neural Architectures—Separating OFDM



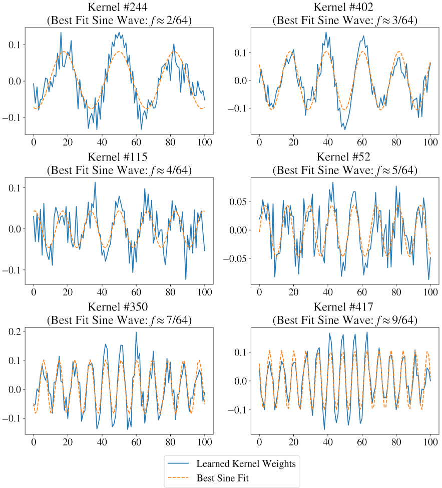
Proposed Modifications: More kernels, longer kernels on 1st convolutional layer[8]

[8] G. Lee, A. Weiss, A. Lancho, Y. Polyanskiy, and G. Wornell, “On Neural Architectures for Deep Learning-Based Source Separation of Co-Channel OFDM Signals,” ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Rhodes Island, Greece, 2023, pp. 1-5.

Effects of Long First-Layer Convolutional Kernel Sizes

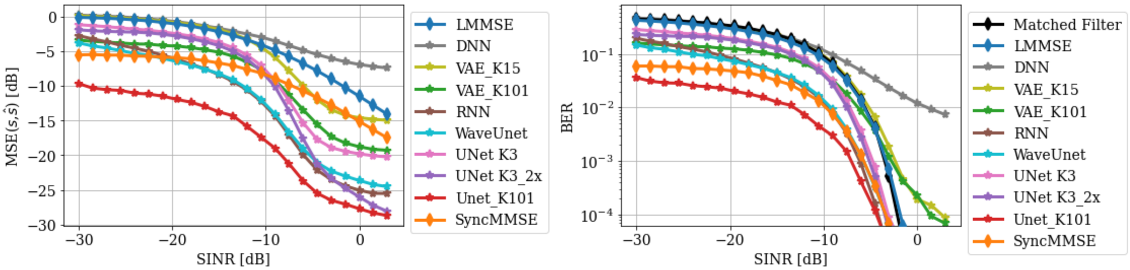


Sharp transition happens
at around true FFT parameter ($K = 64$)



Revisiting Sepration of QPSK SOI + OFDM Interference

- Which neural architecture should I use? Does it matter? Generally YES



Figures from `https://github.com/RFChallenge/SCSS_DNN_Comparison`

Other Data-Driven Approaches

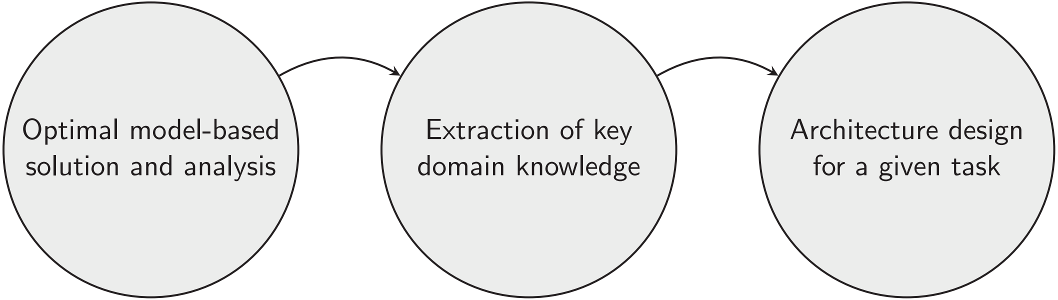
- **End-to-End Separator**
  - Learn a multivariate regression function ("supervised learning")[9]
- **End-to-End Decoder** (for communication SOI)
  - Learn a soft decoder, i.e., output bit probabilities
  - Learn a corresponding decoder from the regression (separator) output
- Learning and using a library of **deep generative priors** *(scalable strategy!)*
  1. Train independent deep generative models for each signal type
  2. Use these models as priors for inference[10]

[9] A. Lancho, A. Weiss, G. Lee, J. Tang, Y. Bu, Y. Polyanskiy, and G. Wornell, "Data-Driven Blind Synchronization and Interference Rejection for Digital Communication Signals," IEEE GLOBECOM 2022, Rio de Janeiro, Brazil, 2022, pp. 2296-2302.

[10] T. Jayakumar, G. Lee, A. Lancho, A. Weiss, Y. Polyanskiy and G. Wornell, "Score-based Source Separation with Applications to Digital Communication Signals," arXiv:2306.14411 [cs.LG], Jun. 2023.

Concluding Remarks

- Systematic development approach to data-driven solutions



- Connecting classical signal processing concepts with architectural choices for data-driven approaches

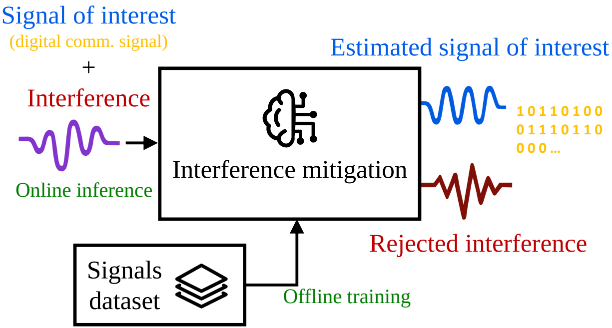
Finding the Appropriate Neural Architecture

- Is a hard problem
  - Appropriate choices lead to improved performance/better generalization with less data
  - Poor choices lead to poor generalization or constraints in optimization space
- Ongoing research to discover effective neural architecture for RF signals
  - **Domain-informed approach**: designing network architectures based on our understanding of the appropriate models
  - Potential automated solution: **Neural Architecture Search**
- Need a **benchmark** to test and compare different neural architectures

  Think MNIST and ImageNet database for image classification/computer vision

MIT RF Challenge[11]

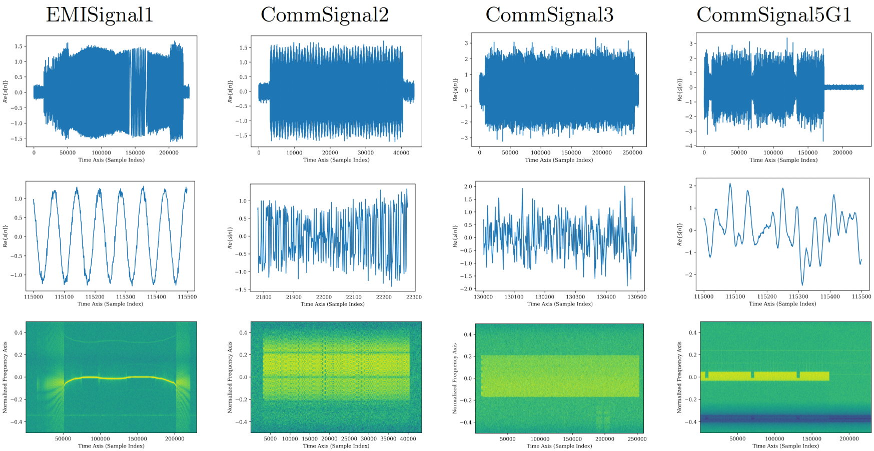
A standardized set of data and tools for the SCSS problem with RF signals



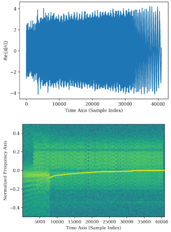
The Challenge: Develop new ML approaches for SCSS and compare performance!

[11] Research was sponsored by the United States Air Force Research Laboratory and the Department of the Air Force Artificial Intelligence Accelerator and was accomplished under Cooperative Agreement Number FA8750-19-2-1000. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Department of the Air Force or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

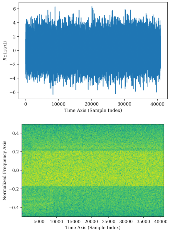
MIT RF Challenge—Data



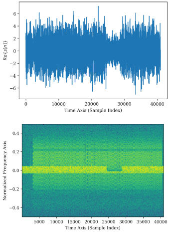
CommSignal2 + EMISignal1



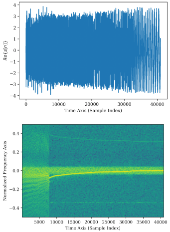
CommSignal2 + CommSignal3



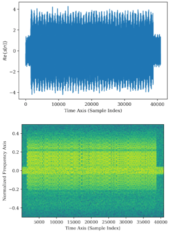
CommSignal2 + CommSignal5G1



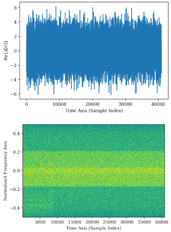
QPSK + EMISignal1



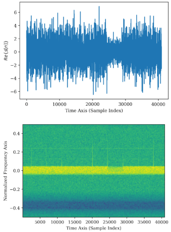
QPSK + CommSignal2



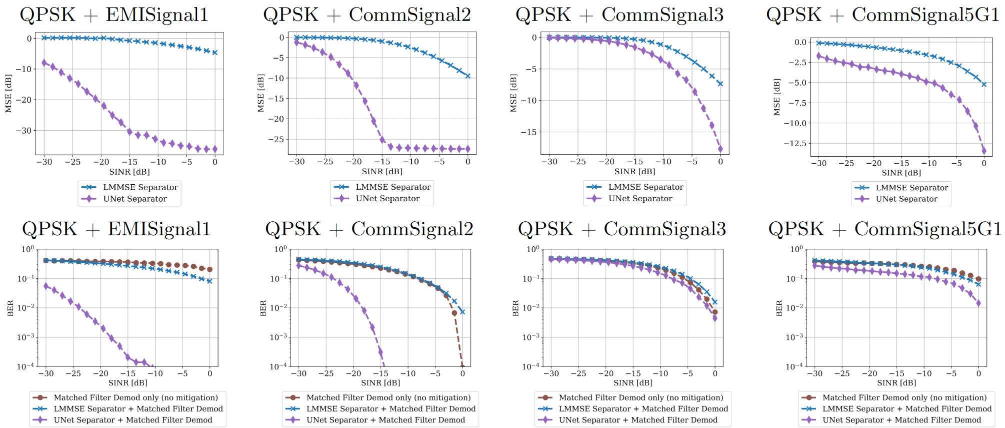
QPSK + CommSignal3



QPSK + CommSignal5G1



MIT RF Challenge—Benchmark Performance



Can you do better?

Hands-On: “Mini” RF Challenge

Community Prediction Competition

Mini RF Challenge

Mini RF Challenge

6 months to go

Overview | Data | Code | Discussion | Leaderboard | Rules | Team | Submissions | **Submit Predictions**

Overview

Start
4 hours ago

Close
6 months to go

Description

As with the proliferation of wireless technologies, radio frequency (RF) bands become crowded, leading to disruptive co-channel interference. The RF Challenge tasks participants with navigating the complexities of separating RF signals amid such interference. Notably, co-channel RF signals present a unique challenge due to overlapping energy content in time and frequency, demanding innovative solutions.

This "Mini RF Challenge" distills the main challenge's essence, allowing participants to delve into a simplified problem set. Participants are to separate a single-carrier QPSK signal-of-interest from interference of unknown signal models, aiming to effectively mitigate the interference and recover the underlying information bits reliably.

While advancements have been made for similar problems in domains like computer vision and audio, the RF realm, with its co-channel signals, remains a challenge. Can you effectively learn the signal structures to facilitate interference mitigation and to improve the demodulation performance of the signal-of-interest?

Competition Host
Gary CF Lee

Prizes & Awards
Kudos

Participation
0 Competitors
0 Teams
0 Entries

Table of Contents