

Estimation, Filtering and Decoding via Deep Learning

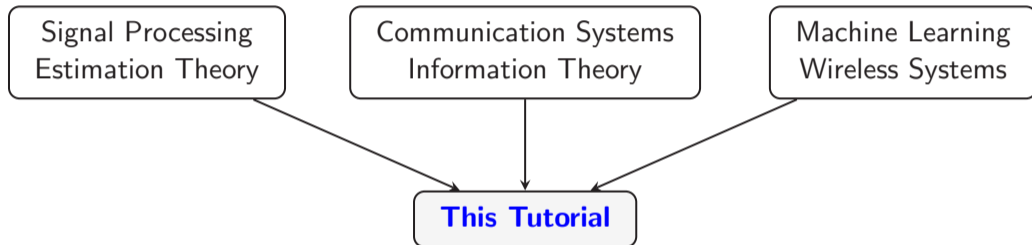
Amir Weiss, Alejandro Lancho and Gary Lee

Massachusetts Institute of Technology, USA

September 4, 2023



- 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

■ What is our motivation? Why are we giving this tutorial?

- Develop this intermediate, hybrid, but so timely and important emerging field



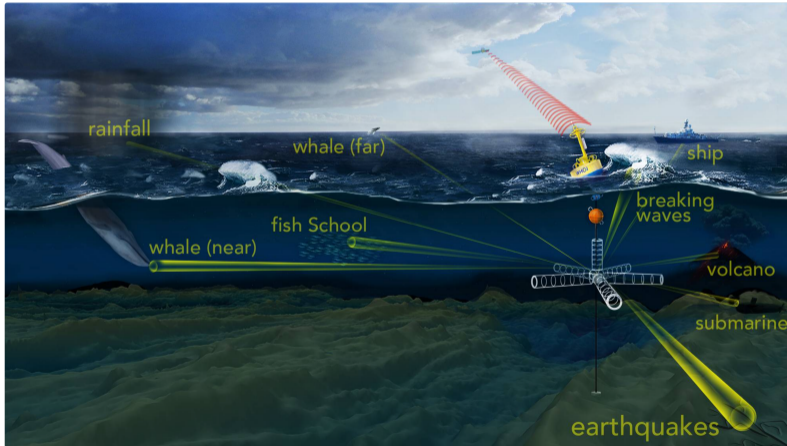
- Sharing important findings, insights and understanding that are not all published
- Present and make easily accessible the RF Challenge

■ 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)

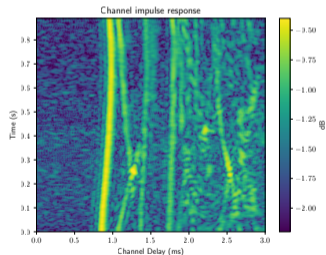


Marine Technology News, March 2021, © Woods Hole Oceanographic Institution, N. Renier

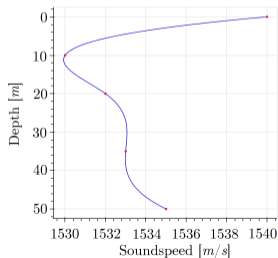
Motivating Applications: Underwater Acoustic Localization

- 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:

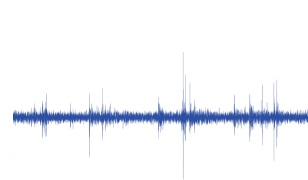
Intricate impulse response



Depth-varying soundspeed profile



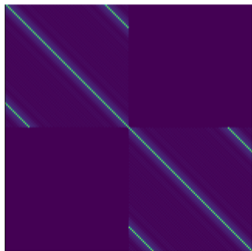
Ambient noise: not Gaussian, not white



Motivating Applications: RF Signal Separation

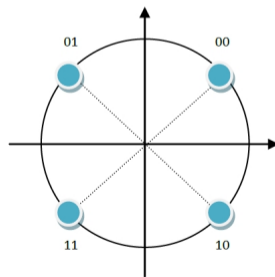
- Increasingly congested spectrum \rightarrow more collisions and overlaps
- **Single-sensor source separation**: key challenge for advanced interference rejection
- **Going beyond stationarity and Gaussianity**, attractive for other problems as well

Nontrivial **temporal structures**



Temporal covariance matrix
of an OFDM signal

Digital communication:
“discrete” nature



■ 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

■ 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

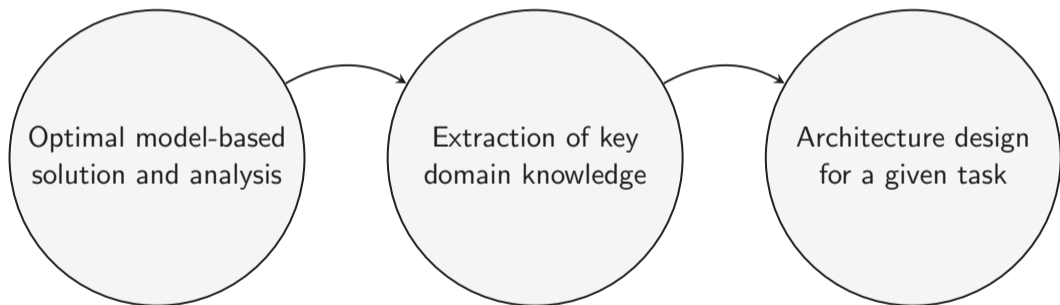
■ Session 3: Deep Learning Methods, Challenges, and a Short Hands-on Session

- On neural architectures for source separation
- RF Challenge/Hands-on Mini RF Challenge

Speaker: Gary Lee

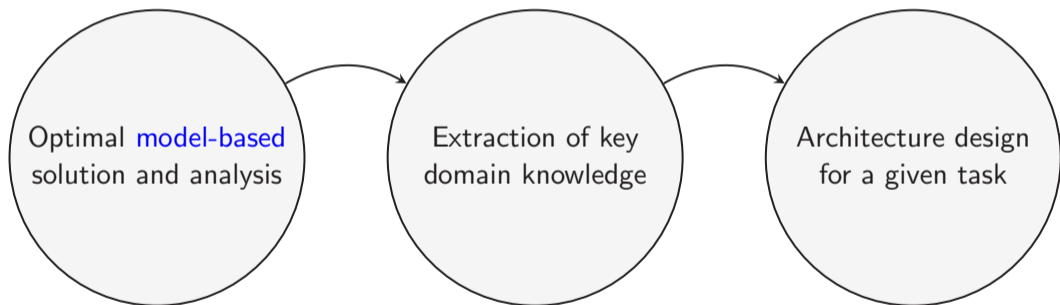
Balancing The “Model-Based” and “Data-Driven” Approaches

- 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

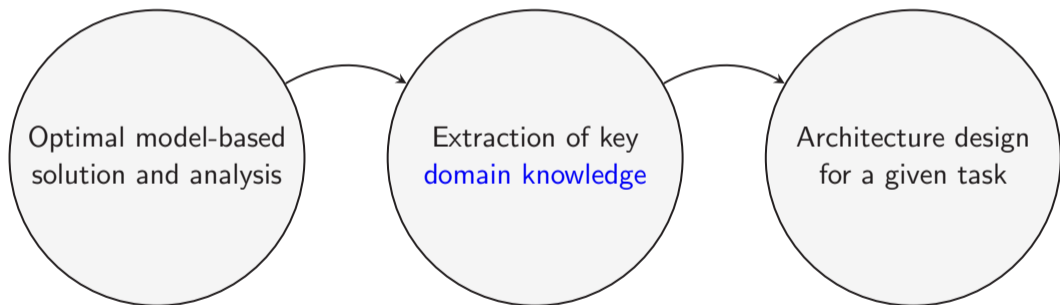
- We view the two approaches, rather than contrasting, as **complementary**
- Design process of an ML-aided solution for a given problem:



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

Balancing The “Model-Based” and “Data-Driven” Approaches

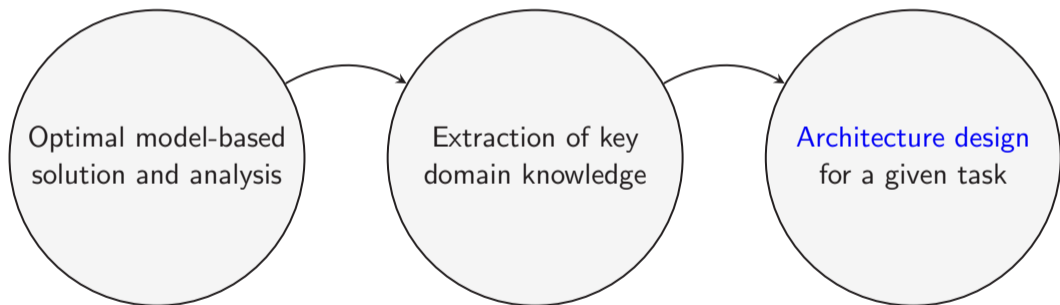
- We view the two approaches, rather than contrasting, as **complementary**
- Design process of an ML-aided solution for a given problem:



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:

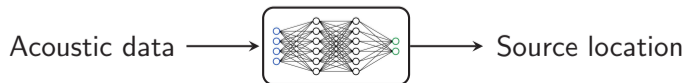


Input structure, key design parameters, training procedure, loss function(s)

Estimation: Localization as a Case Study

- Underwater localization: Enabling technological ability for a variety of applications
- Acoustic waves → favorable propagation properties underwater
- Physics of underwater acoustics is (relatively) well-understood
- **However**, analytically complicated → 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?



Data-driven direct localization with **single-sensor receivers**

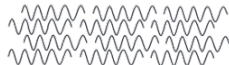
- Model is generally unknown, **availability of datasets**



- **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:

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

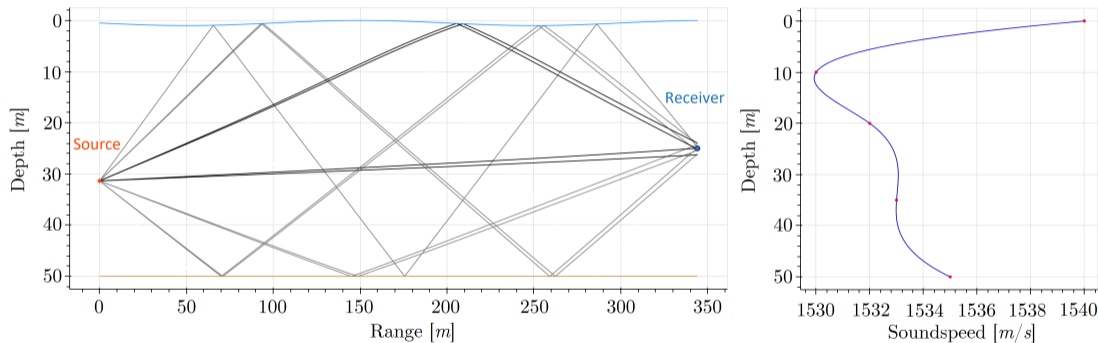
- \mathbf{x}_ℓ : received signal at the ℓ -th receiver (**observed**)
- \mathbf{s} : emitted waveform from the acoustic source (**unknown**)
- \mathbf{v}_ℓ : additive noise, not necessarily Gaussian/white
- $\mathbf{H}_\ell(\mathbf{p}, \mathcal{E}) = \text{Diag}(\mathbf{h}(\mathbf{p}, \mathcal{E}))$: 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}

The UWA Localization Problem

- Illustration of simulated ray propagation model in nonisovelocity environment

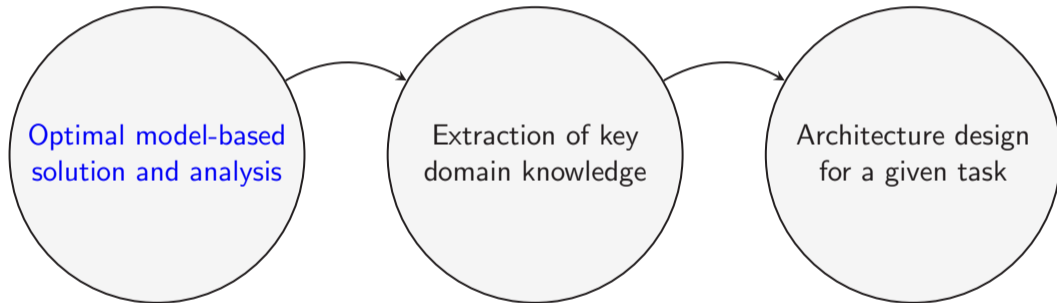


Plots generated using the Bellhop simulator

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

Re-Our Proposed Methodology

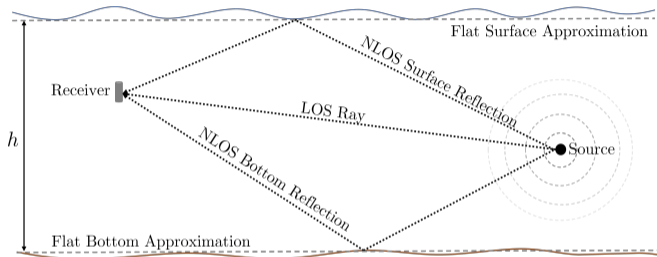
- Design process of an ML-aided solution for a given problem:



“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

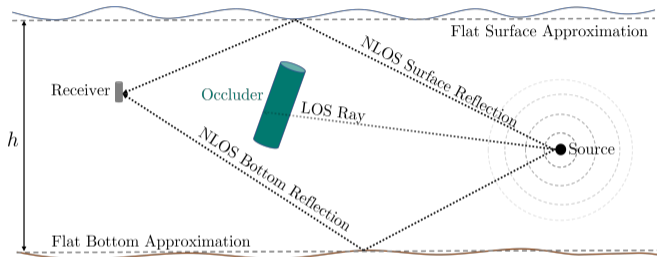


- $$[\mathbf{H}_\ell(\mathbf{p}, \mathcal{E})]_{kk} = \underbrace{b_{1\ell} e^{-j\omega_k \tau_{1\ell}(\mathbf{p}, \mathcal{E})}}_{\text{LOS}} + \underbrace{b_{2\ell} e^{-j\omega_k \tau_{2\ell}(\mathbf{p}, \mathcal{E})}}_{\text{NLOS, surface}} + \underbrace{b_{3\ell} e^{-j\omega_k \tau_{3\ell}(\mathbf{p},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

¹Weiss, A., Arıkan, 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



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

- For example, one special case: $b_{1\ell} = 0$

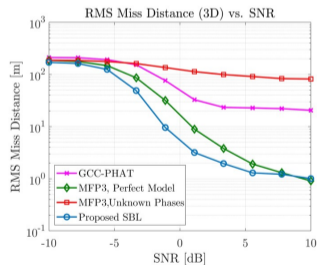
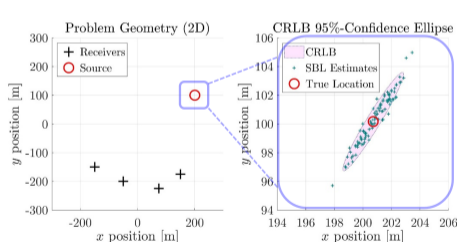
¹Weiss, A., Arıkan, T., Vishnu, H., Deane, G.B., Singer, A.C. and Wornell, G.W., 2022. A semi-blind method for localization of underwater acoustic sources. IEEE Transactions on Signal Processing, 70, pp.3090-3106.

Step 1: A Model-Based Optimal Solution

- Optimal solution¹ (in the least-squares sense):

$$\hat{\mathbf{p}}_{\text{SBL}} = \arg \max_{\mathbf{p} \in \mathbb{R}^{3 \times 1}} \lambda_{\max} \left(\underbrace{\mathbf{Q}(\mathbf{p}, \mathbf{x}_1, \dots, \mathbf{x}_L)}_{\text{Position- and data-dependent matrix}} \right)$$

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



¹Weiss, A., Arıkan, 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 solution¹ (in the least-squares sense):

$$\hat{\mathbf{p}}_{\text{SBL}} = \arg \max_{\mathbf{p} \in \mathbb{R}^{3 \times 1}} \lambda_{\max} \left(\underbrace{\mathbf{Q}(\mathbf{p}, \mathbf{x}_1, \dots, \mathbf{x}_L)}_{\text{Position- and data-dependent matrix}} \right)$$

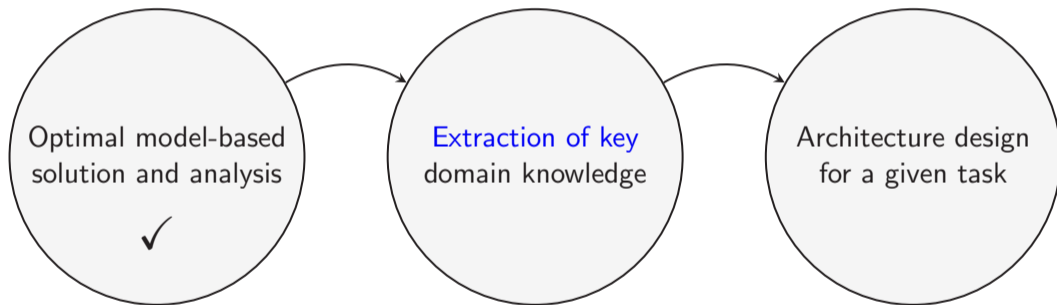
- In white Gaussian noise, for spectrally flat signal, attains the Cramér-Rao lower bound
 - Trivially extends to an R -ray model with $R > 3$
 - Extends to an nonisovelocity propagation model
 - Extends to nonflat ocean surface and bottom

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

¹Weiss, A., Arıkan, 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

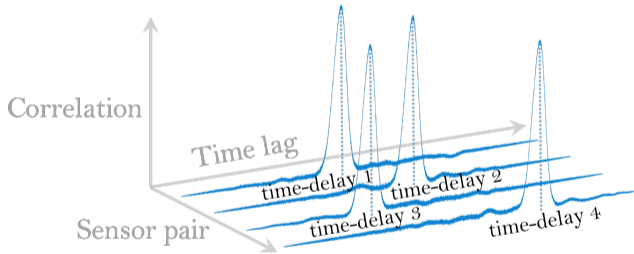
- **Inference computational complexity:** can be reduced?

$$\hat{\mathbf{p}}_{\text{SBL}} = \arg \max_{\mathbf{p} \in \mathbb{R}^{3 \times 1}} \lambda_{\max} \left(\underbrace{\mathbf{Q}(\mathbf{p}, \mathbf{x}_1, \dots, \mathbf{x}_L)}_{\text{Position- and data-dependent matrix}} \right)$$

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**,

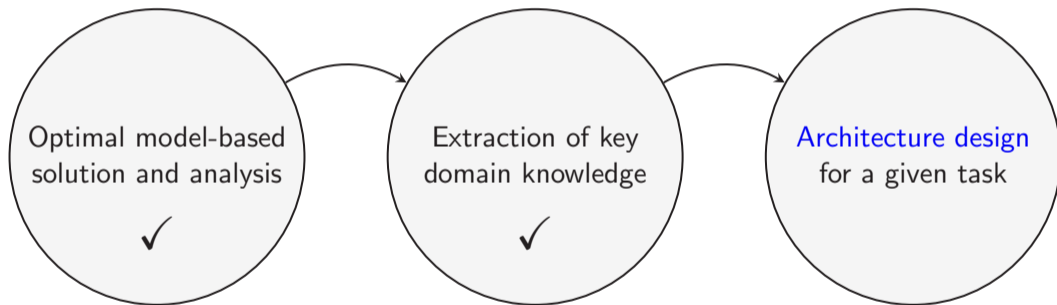
Objective

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

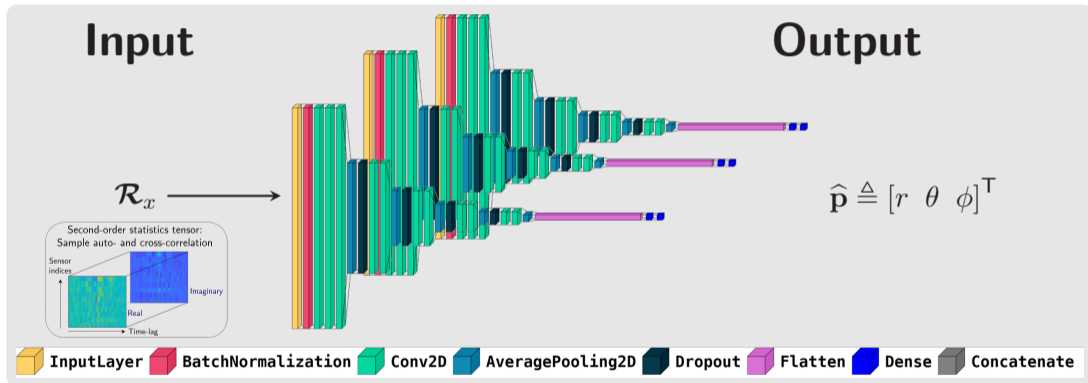
- Design process of an ML-aided solution for a given problem:



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



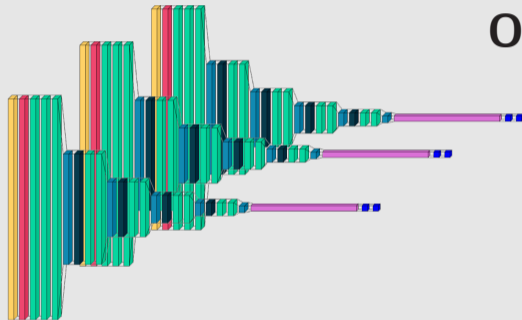
The model is comprised of **three pre-trained sub-models**

The Proposed Solution

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

Input

\mathcal{R}_x →



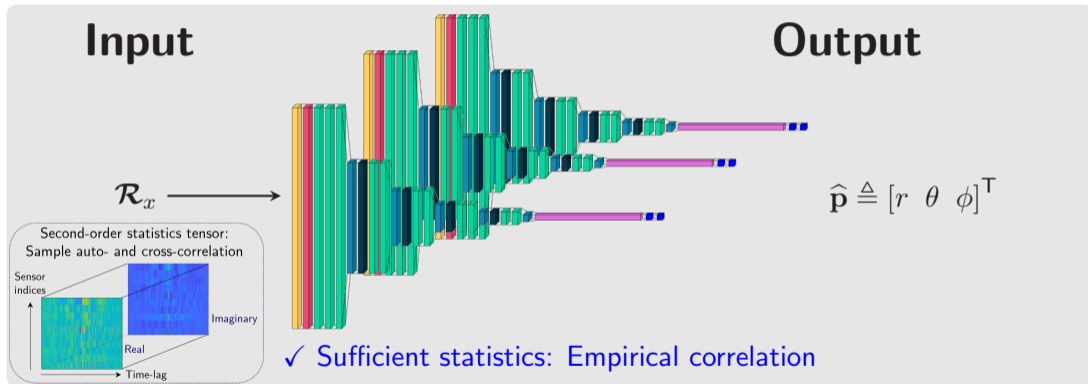
Output

$$\hat{\mathbf{p}} \triangleq [r \ \theta \ \phi]^T$$

✓ Overall computational complexity at inference time: fixed

The Proposed Solution

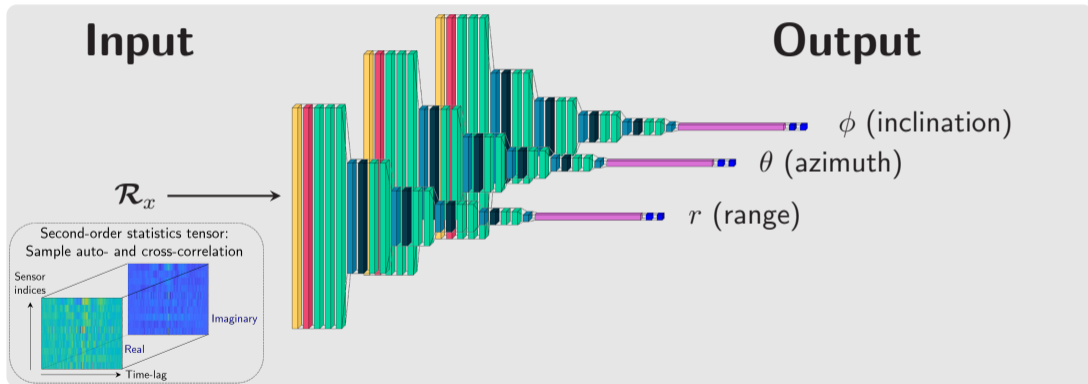
- If we use a NN for the solution, what **input structure** should it have?



Conv2D layers + Long kernel size at the first layer

The Proposed Solution

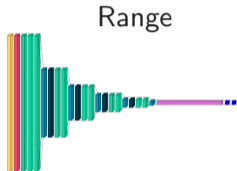
- Exploit statistical dependencies between, e.g., azimuth and range?



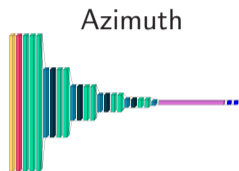
✓ The model is comprised of **three pre-trained sub-models**

Progressive Training and Loss Functions

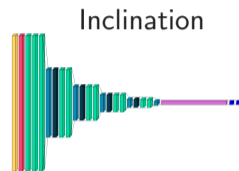
■ Phase 1: Train individual models



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



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



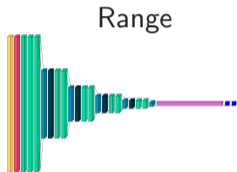
$$\mathcal{L}_\varphi \triangleq 2 - 2 \cos(2(\hat{\varphi}(\mathbf{w}_\varphi) - \varphi))$$

Empirical cyclic error

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

Progressive Training and Loss Functions

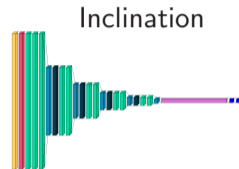
■ Phase 1: Train individual models



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

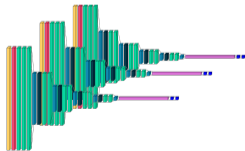


$$\mathcal{L}_\theta \triangleq 2 - 2 \cos(\hat{\theta}(\mathbf{w}_\theta) - \theta)$$



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

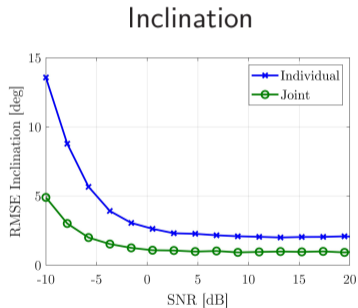
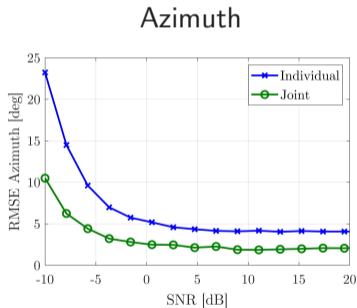
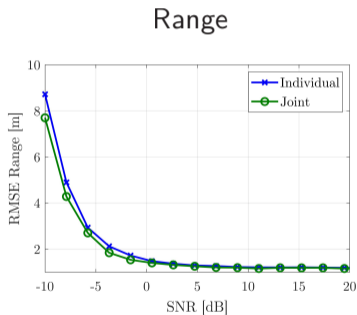
■ Phase 2: Train global model with “hot” initialization (joint optimization approximator)



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

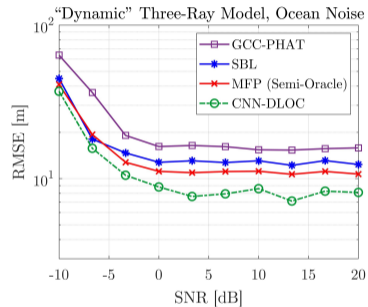
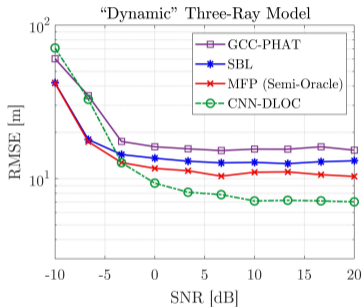
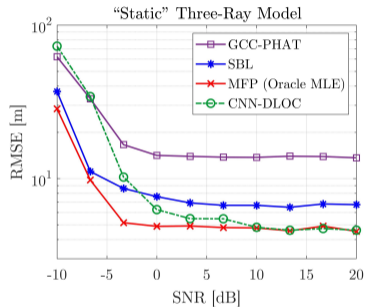
Simulation Results

- 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 $50m$



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

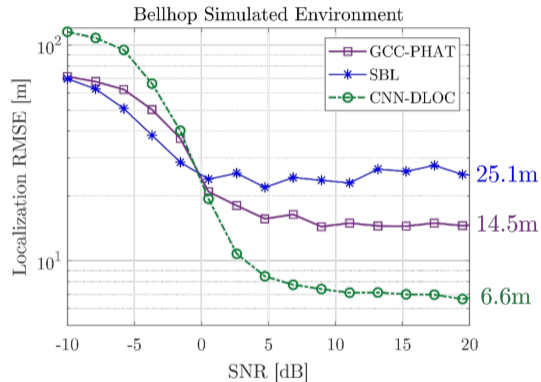
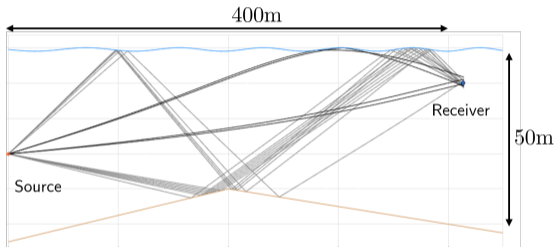
Simulation Results: 3-Ray Model



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

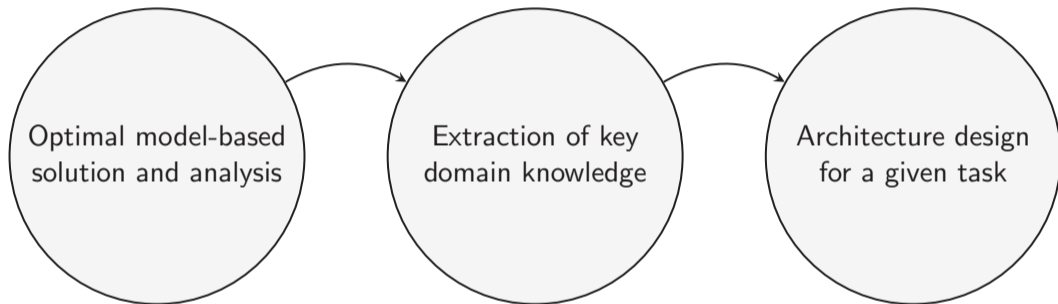
Simulation Results: Bellhop Simulated Environment



- **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**)

■ 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

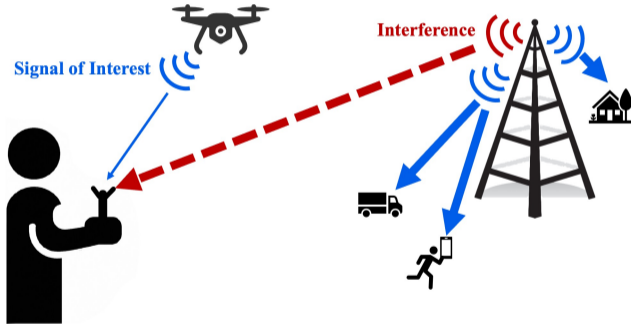
■ Session 3: Deep Learning Methods, Challenges, and a Short Hands-on Session

- On neural architectures for source separation
- RF Challenge/Hands-on Mini RF Challenge

Speaker: Gary Lee

Motivation

- Radio spectrum increasingly crowded → **spectrum sharing unavoidable**
 - To keep **high reliabilities** → **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**)

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

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

\rightarrow statistically independent

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

- For a recording of N samples:

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

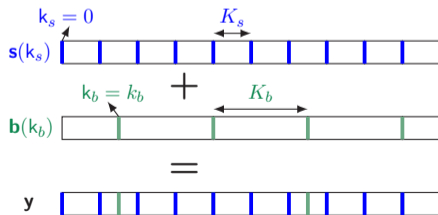
Problem Setup II

$$\mathbf{y} = \mathbf{s}(k_s) + \rho_{\text{SIR}}^{-1/2} \mathbf{b}(k_b) + \rho_{\text{SNR}}^{-1/2} \mathbf{w}$$

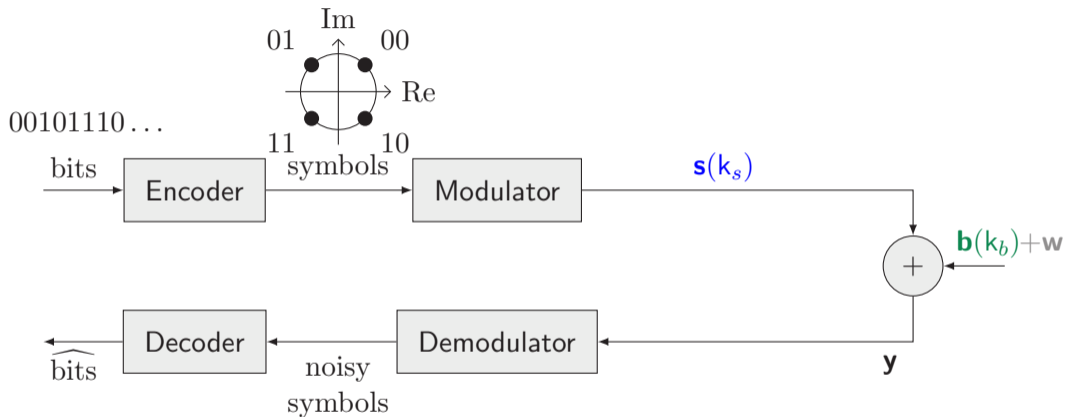
■ $\mathbf{s}(k_s), \mathbf{b}(k_b)$ assumed to be zero-mean, unit-variance, **cyclostationary signals**

- K_s, K_b : cyclic periods
- k_s, k_b : **arbitrary time shifts** w.r.t. start of cyclic period of $\mathbf{s}[n], \mathbf{b}[n]$

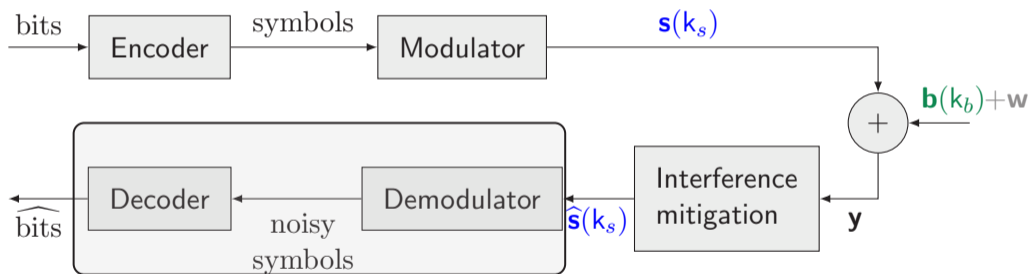
$$\rightarrow k_s \sim \text{Unif}\{1, \dots, K_s\}, k_b \sim \text{Unif}\{1, \dots, K_b\}$$



Scheme of a Digital Communication System

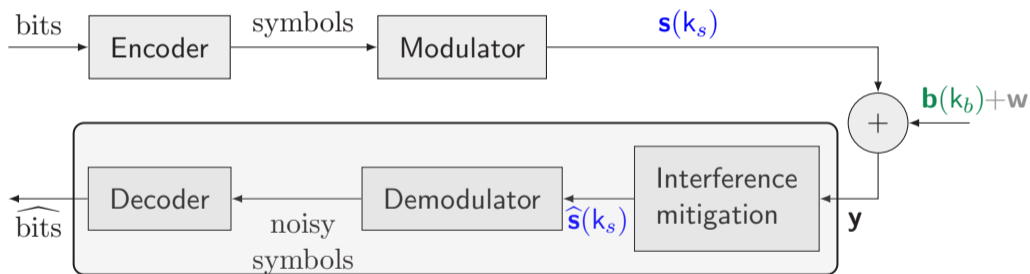


Scheme of a Digital Communication System



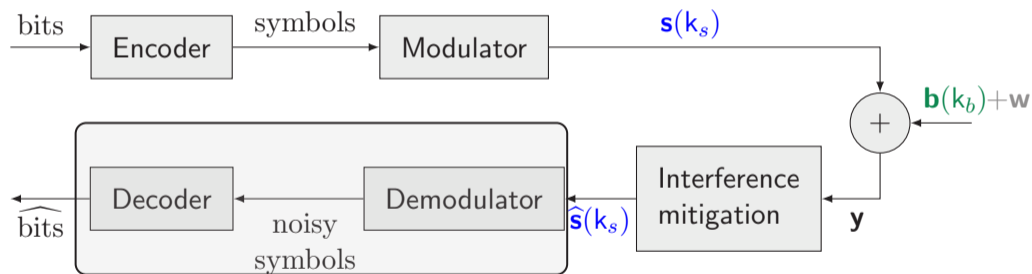
separation + demodulation

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**):

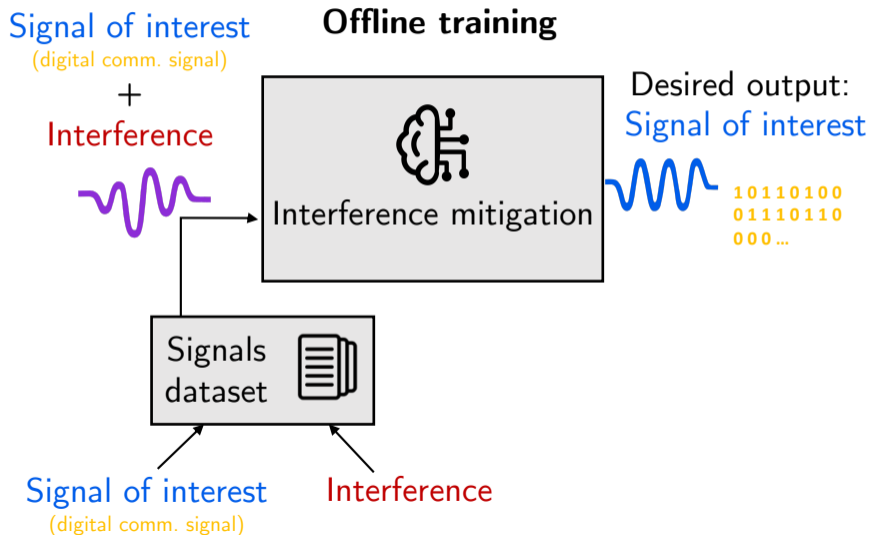
$$\hat{\mathbf{s}} = \mathbf{C}_{ss}(\mathbf{C}_{ss} + \mathbf{C}_{vv})^{-1} \mathbf{y}$$

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

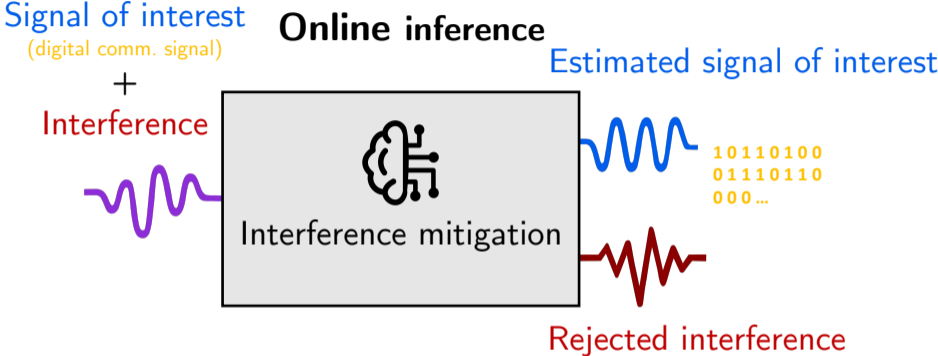
■ Potential problem: They can be applied in different (“small”) time scales

- **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

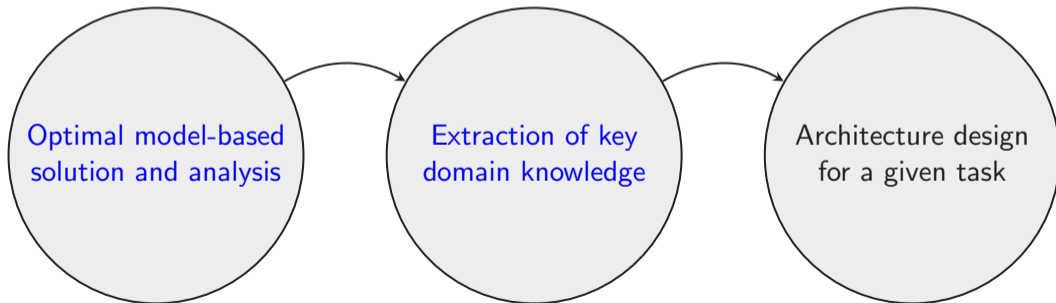
- Deep neural networks (DNNs) successful for source separation
 - **Computer vision**: Color features and local features
 - **Audio**: Spectrogram masking methods

Many communication signals

- May not have local features/dependencies (e.g. OFDM signals)
 - Overlapping in time and frequency
- ⇒ **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**
 - From model-based to data-driven solutions:



A Signal Model Beyond Stationarity

■ Cyclostationary Gaussian mixture model:

$$\mathbf{y} = \mathbf{s}(k_s) + \rho_{\text{SIR}}^{-1/2} \mathbf{b}(k_b) + \rho_{\text{SNR}}^{-1/2} \mathbf{w} = \mathbf{s}(k_s) + \mathbf{v}(k_b)$$

- $\mathbf{s}(m_s) \sim \mathcal{CN}(\mathbf{0}, \mathbf{C}_{ss}(m_s))$, $\mathbf{b}(m_b) \sim \mathcal{CN}(\mathbf{0}, \mathbf{C}_{bb}(m_b))$, $\mathbf{w} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$
- K_s, K_b : cyclic periods
- k_s, k_b : **arbitrary time shifts** w.r.t. start of cyclic period of $\mathbf{s}[n]$, $\mathbf{b}[n]$
→ $k_s \sim \text{Unif}\{1, \dots, K_s\}$, $k_b \sim \text{Unif}\{1, \dots, K_b\}$
- $\rho_{\text{SIR}}, \rho_{\text{SNR}} \in \mathbb{R}_+$ → **assumed to be known/ can be estimated**

Objective

Obtain understanding based on analysis

⇒ **Make informed architectural decisions**

Assume Models of $\mathbf{s}(k_s)$ and $\mathbf{b}(k_b)$ Are Known

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

$$\begin{aligned}\hat{\mathbf{s}}_{\text{MMSE}} &= \mathbb{E} [\mathbb{E}[\mathbf{s}(k_s) | \mathbf{y}, k_s, k_b] | \mathbf{y}] = \mathbb{E} [\hat{\mathbf{s}}_{\text{CLMMSE}}(k_s, k_b) | \mathbf{y}] \\ &= \sum_{m_s=1}^{K_s} \sum_{m_b=1}^{K_b} \mathbb{P}[k_s = m_s, k_b = m_b | \mathbf{y}] \hat{\mathbf{s}}_{\text{CLMMSE}}(m_s, m_b)\end{aligned}$$

with

$$\begin{aligned}\hat{\mathbf{s}}_{\text{CLMMSE}}(m_s, m_b) &\triangleq \mathbf{C}_{sy}(m_s, m_b) \mathbf{C}_{yy}(m_s, m_b)^{-1} \mathbf{y} \\ &= \mathbf{C}_{ss}(m_s) (\mathbf{C}_{ss}(m_s) + \mathbf{C}_{vv}(m_b))^{-1} \mathbf{y}\end{aligned}$$

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

Is the Previous Assumption Reasonable?

■ Three main problems:

- **Computing** $\mathbb{P}[\mathbf{k}_s = m_s, \mathbf{k}_b = m_b | \mathbf{y}] \forall m_s, m_b$ 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 synchronized dataset
- **Two step synchronization-separation:**
 - **MAP estimation of time shift:** $\hat{k}_b^{\text{MAP}} \triangleq \arg \max_{m \in \{1, \dots, K_b\}} \mathbb{P}[k_b = m | \mathbf{y}]$
 - **MAP-based quasi-LMMSE estimator:** $\hat{\mathbf{s}}_{\text{MAP-QLMMSE}} \triangleq \hat{\mathbf{s}}_{\text{LMMSE}}(\hat{k}_b^{\text{MAP}})$
- We show that the MAP-QLMMSE estimator is **asymptotically optimal**
 - Under mild condition (shift uniquely detectable):

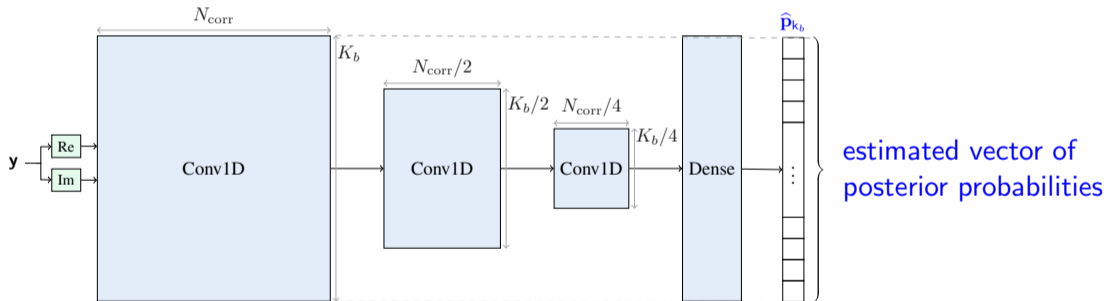
$$\mathbb{P}[\hat{k}_b^{\text{MAP}} \neq k_b] = o\left(\frac{1}{N^\alpha}\right), \quad \lim_{N \rightarrow \infty} \frac{\mathbb{E}[\|\hat{\mathbf{s}}_{\text{LMMSE}} - \mathbf{s}\|_2^2]}{\mathbb{E}[\|\hat{\mathbf{s}}_{\text{MAP-QLMMSE}} - \mathbf{s}\|_2^2]} = 1$$

■ MAP estimator computationally hard

²A. 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

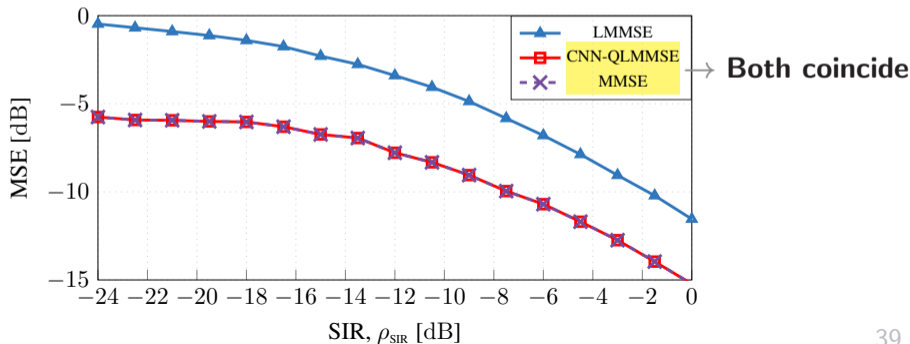
- Assume **availability of dataset** $\{\mathbf{s}(k_s = 0)^{(i)}, \mathbf{b}(k_b = k_b)^{(i)}\}, i \in \{1, \dots, T\}$
 \Rightarrow **Data-driven approach** to estimate \hat{k}_b^{MAP} via CNNs \rightarrow **CNN-QLMMSE**



Separation Performance on Short Gaussian Mixtures

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

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

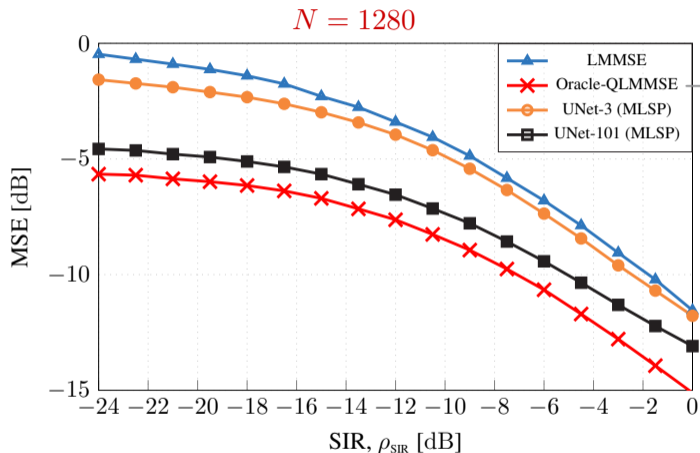


Revisiting the Assumptions Made

■ Three main problems:

- **Computing** $\mathbb{P}[\mathbf{k}_s = m_s, \mathbf{k}_b = m_b | \mathbf{y}] \forall m_s, m_b$ 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**

Simulation Results on Short Gaussian Mixture³



Requires oracle access to synchronized statistics and correct shift

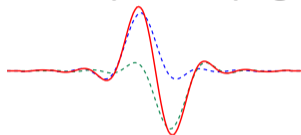
³G.C.F. Lee, A. Weiss, A. Lancho, J. Tang, Y. Bu, Y. Polyanskiy and G.W. Wornell, "Exploiting temporal structures of cyclostationary signals for data-driven single-channel source separation," in Proc. IEEE International Workshop for Machine Learning and Signal Processing (MLSP), Xi'an, China, Aug. 2022.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

Beyond Gaussianity: Digital Communication Signals

$$\mathbf{y} = \mathbf{s}(k_s = 0) + \rho_{\text{SIR}}^{-1/2} \mathbf{b}(k_b) + \rho_{\text{SNR}}^{-1/2} \mathbf{w}$$

- $\mathbf{s}(k_s = 0)$ bears **QPSK symbols** using a **root-raised cosine pulse-shaping filter**

- Spanning 8 symbols, oversampling factor = 16



- $\mathbf{b}(k_b)$ bears **16QAM OFDM symbols** of length $K_b = 80$ with $k_b \sim \text{Unif}\{1, \dots, K_b\}$

- FFT size = 64, cyclic-prefix length = 16

- Details on signals generation process → **Visit our Github repository:**

https://github.com/RFChallenge/SCSS_Sync

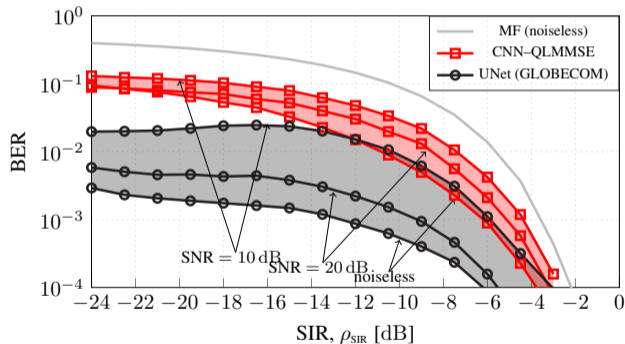
- **Performance metric:** BER

- Every approach includes a **last (standard) MF step prior to hard decoding**

Simulation Results for Different SNR Values⁴

■ Input length:

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



⁴A. 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.

- Learned that from the **cyclostationary Gaussian mixture model**
 - **Synchronization** (dataset level or shifts) → significant **performance gains**
 - **End-to-end DNN** architecture needs to be able to **synchronize** (even implicitly)
 - **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?
 - **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):
<https://bit.ly/RFHandsOn2023>)*

■ 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

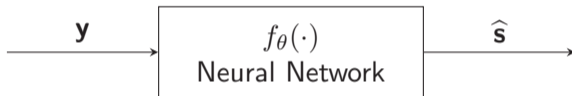
■ Session 3: Deep Learning Methods, Challenges, and a Short Hands-on Session

- On neural architectures for source separation
- RF Challenge/Hands-on Mini RF Challenge

Speaker: Gary Lee

SCSS as a Multivariate Regression Problem

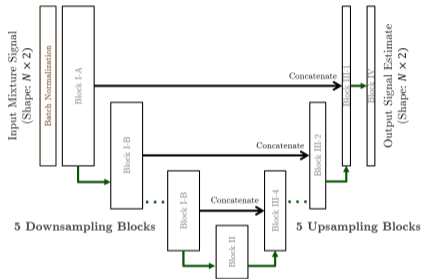
- Learning an end-to-end separator
- Appropriate Parameterization \Rightarrow Neural Architecture Choice



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

Minimum Mean-Square Error Estimator

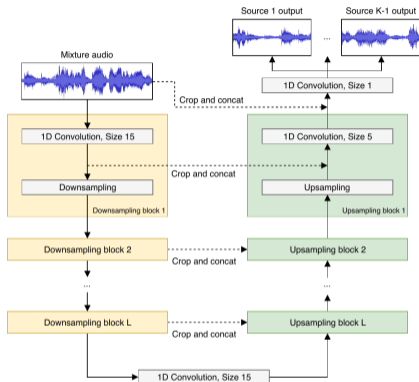
U-Net Architecture



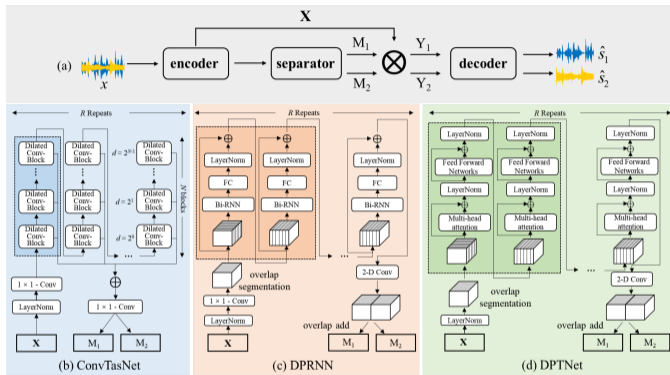
- A fully convolutional network architecture with the same input and output size
- First used in biomedical image segmentation⁵
- Successive downsampling and upsampling blocks (multiresolution features)

⁵O. 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-Net⁶



TasNet⁷

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

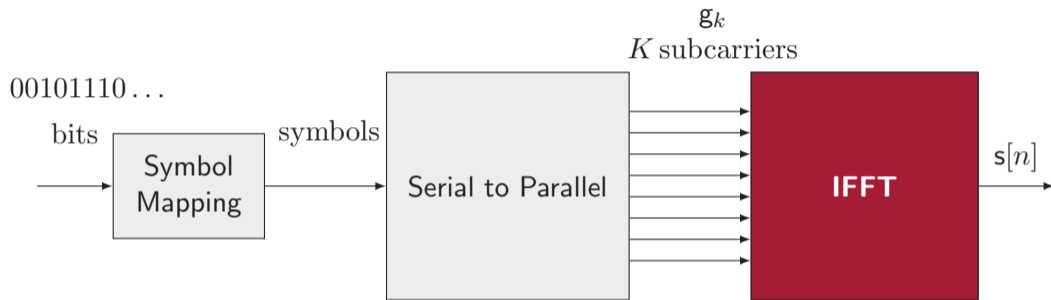
- Which neural architecture should I use? Does it matter?
 - Appropriate architecture leads to more efficient training/better generalization
 - Architectures are typically chosen based on precedence, intuition, and trial and error
 - *What works for image/audio might not work for RF signals (?)*
- Demonstration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



$$s[n] = \sum_{k=0}^{K-1} g_k \exp(j2\pi kn/K)$$

Comparing Neural Architectures—Separating OFDM

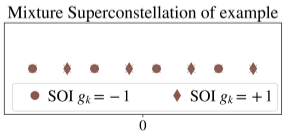
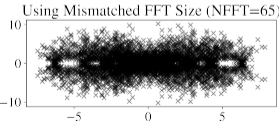
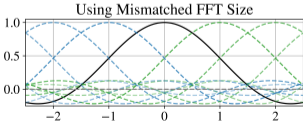
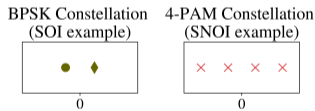
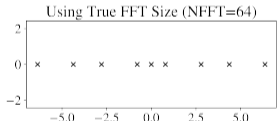
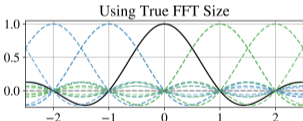
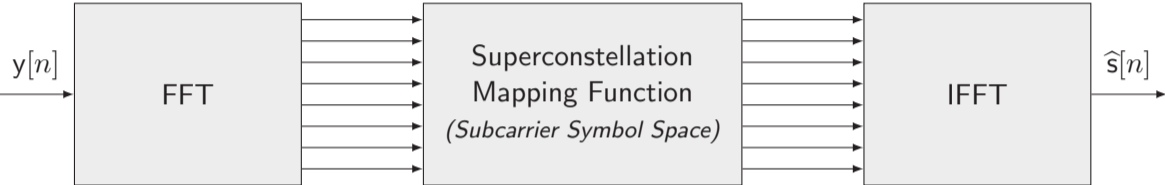
(Unobserved)

$$\begin{aligned} \mathbf{s}[n] &= \sum_{k=0}^{K-1} \mathbf{g}_k r[n - T_{\text{cp}}, k], \quad \mathbf{b}[n] = \sum_{k=0}^{K-1} \mathbf{h}_k r[n - T_{\text{cp}}, k], \\ r[n, k] &\triangleq \exp(j2\pi kn/K) \mathbb{1}_{\{-T_{\text{cp}} \leq n < K\}}, \end{aligned}$$

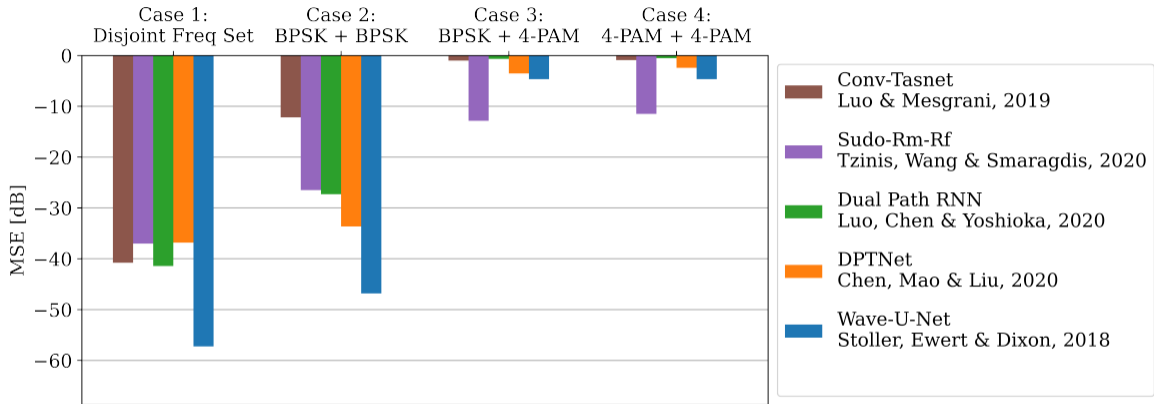
$$\begin{aligned} \mathbf{y}[n] = \mathbf{s}[n] + \mathbf{b}[n] &= \sum_{k=0}^{K-1} \mathbf{a}_k \exp(j2\pi k(n - T_{\text{cp}})/K) \mathbb{1}_{\{0 \leq n < K + T_{\text{cp}}\}}, \\ \mathbf{a}_k &= \mathbf{g}_k + \mathbf{h}_k, \quad \mathbf{a}_k \in \mathcal{A}. \end{aligned}$$

Goal: Estimate \mathbf{s} from observation \mathbf{y} .

Separating OFDM: (Oracle) Model-Based Approach

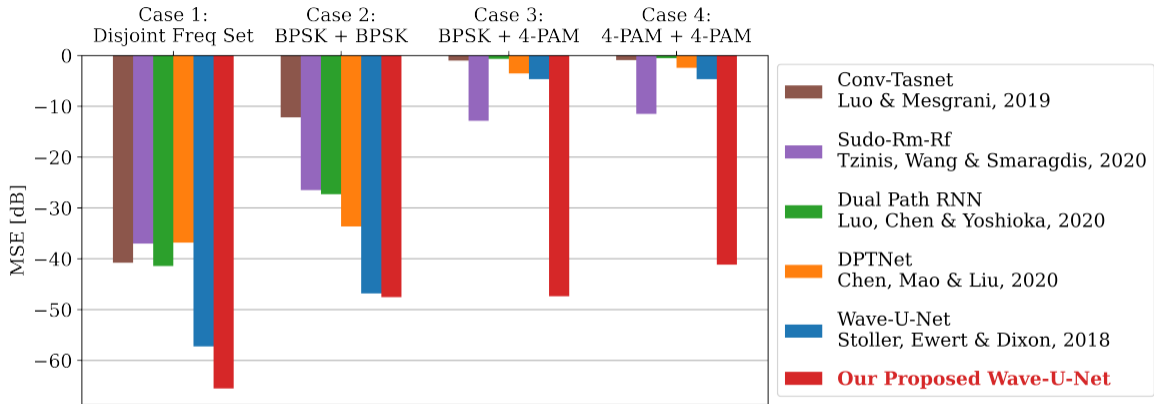


Comparing Neural Architectures—Separating OFDM



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

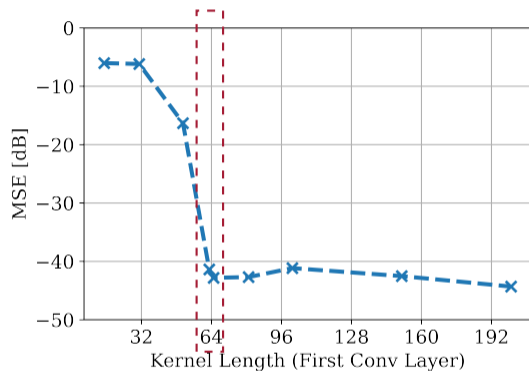
Comparing Neural Architectures—Separating OFDM



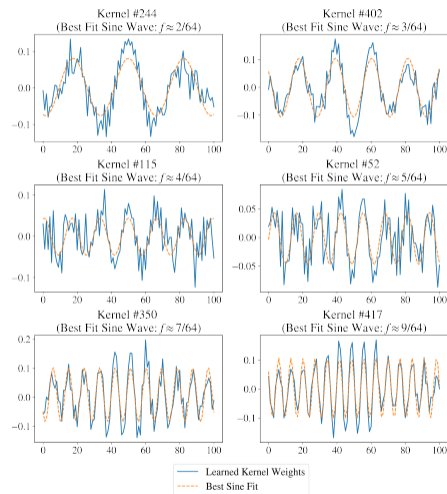
Proposed Modifications: More kernels, longer kernels on 1st convolutional layer⁸

⁸G. 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

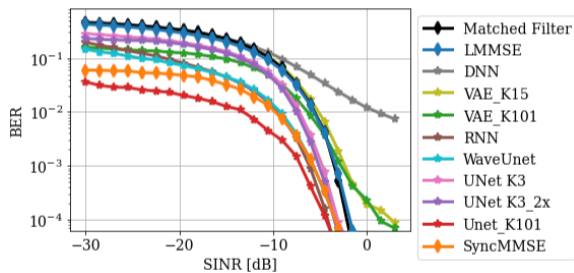
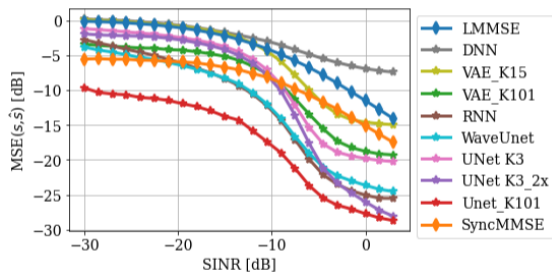


Sharp transition happens
at around true FFT parameter ($K = 64$)



Revisiting Separation 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

■ 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!*)

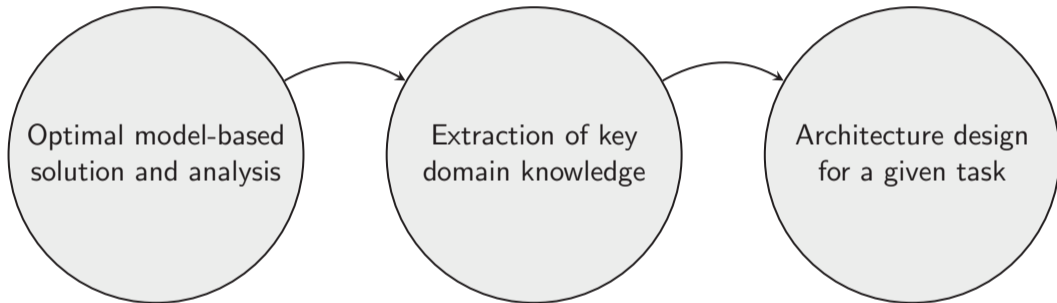
1. Train independent deep generative models for each signal type
2. Use these models as priors for inference¹⁰

⁹A. 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.

¹⁰T. 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



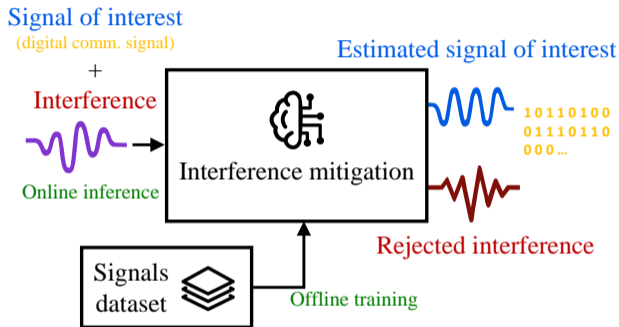
- Connecting classical **signal processing concepts** with **architectural choices** for data-driven approaches

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

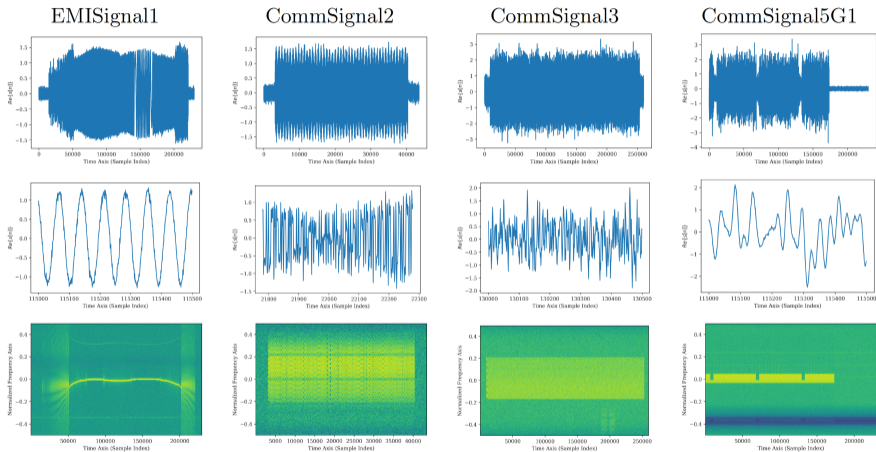
A standardized set of data and tools for the SCSS problem with RF signals



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

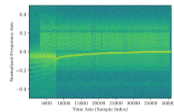
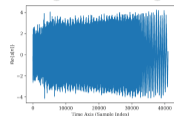
¹¹Research was sponsored by the United States Air Force Research Laboratory and the Department of the Air Force Artificial Intelligence Accelerator and was accomplished under Cooperative Agreement Number FA8750-19-2-1000. The views and conclusions contained in this document are those of the authors and 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

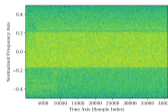
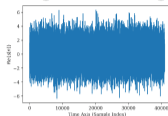


MIT RF Challenge—Separation and Demodulation Sub-Challenges

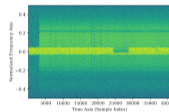
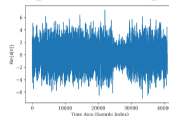
CommSignal2 + EMISignal1



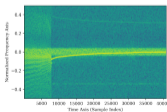
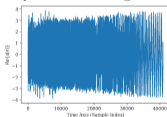
CommSignal2 + CommSignal3



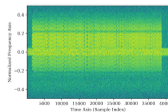
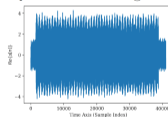
CommSignal2 + CommSignal5G1



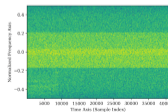
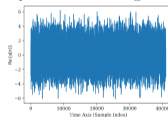
QPSK + EMISignal1



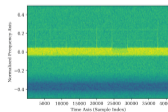
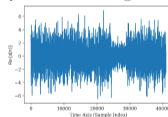
QPSK + CommSignal2



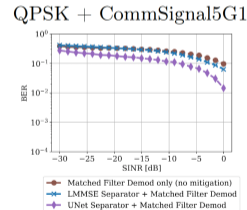
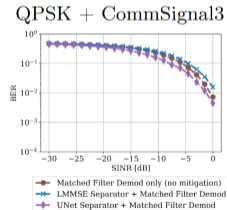
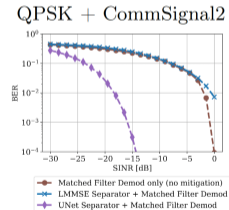
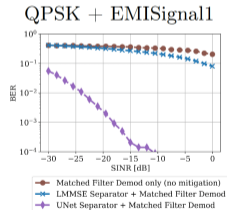
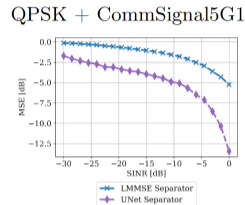
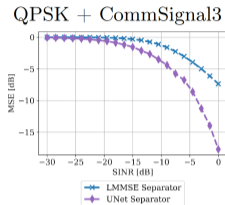
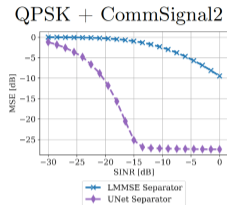
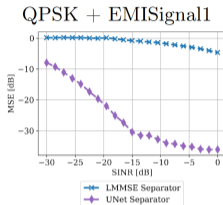
QPSK + CommSignal3



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

Description

Evaluation

Citation