Learning the Barankin Lower Bound on DOA Estimation Error

Hai Victor Habi[1] Hagit Meseer[1] and Yoram Bresler[2]

[1]School of Electrical Engineering, Tel Aviv University, Israel
[2]Department of ECE, University of Illinois, Urbana-Champaign, USA

Hagit Messer

Yoram Bresler

1. Problem Statement
2. Introducing the concept of a Generative Barankin Bound (GBB) - a learned performance bound for understanding large errors in direction of arrival problems.
3. Suggesting *DOA Flow* – A Conditional Normalizing Flow Model for Direction of Arrival Estimation.
4. Experimental results for the GBB on DOA estimation errors for cases that are analytically intractable.
5. Conclusions.

Barankin Bound

Let,

- $\boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}$ be a parameter.
- $\boldsymbol{X} \in \Upsilon \subseteq \mathbb{R}^{d_x}$ be a measurement vector.
- $\boldsymbol{\psi} \in \Theta$ be a test point.

Then for any unbiased estimator $\hat{\boldsymbol{\theta}}(\cdot)$ we have:

$$\mathrm{Cov}\left(\hat{\boldsymbol{\theta}}\right) \triangleq \mathbb{E}_{\boldsymbol{X}}\left[\left(\hat{\boldsymbol{\theta}}(\boldsymbol{X}) - \boldsymbol{\theta}\right)^2\right] \geq \mathrm{BB} \triangleq \frac{\Delta^2}{b(\Delta) - 1}, \quad (1)$$

$$b(\Delta) \triangleq \mathbb{E}_{\boldsymbol{X}}\left[\eta^2(\boldsymbol{X}; \boldsymbol{\theta}, \Delta)\right], \quad (2)$$

where,

- $\Delta = \boldsymbol{\psi} - \boldsymbol{\theta}$ is the deviation of the test point from parameter $\boldsymbol{\theta}$.
- $\eta(\boldsymbol{x}; \boldsymbol{\theta}, \Delta) \triangleq \frac{f_{\boldsymbol{X}}(\boldsymbol{x}; \Delta + \boldsymbol{\theta})}{f_{\boldsymbol{X}}(\boldsymbol{x}; \boldsymbol{\theta})}$ is the likelihood ratio.

Our goal is to study the threshold effect (using the Barankin bound) when $f_{\boldsymbol{X}}(\boldsymbol{x}; \boldsymbol{\theta})$ is completely unknown but a data set $\mathcal{D}_T = \{\boldsymbol{x}_i, \boldsymbol{\theta}_i\}_{i=1}^n$ of samples is available.

Approach

Stage 1: Learn a generative model of the true measurement distribution $f_{\boldsymbol{X}}(\boldsymbol{x};\boldsymbol{\theta})$ using the dataset $\mathcal{D}_T=\{\boldsymbol{x}_i,\boldsymbol{\theta}_i\}_{i=1}^{N_v}$ of N_v measurement-parameter sample pairs.
Stage 2: Use the learned generative model to approximate the Barankin bound.

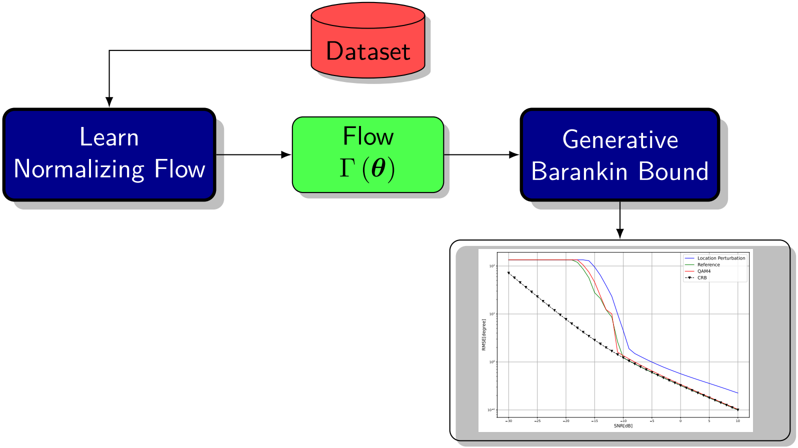

Figure: Approach Overview

Approach

Stage 1: Learn a conditional normalizing flow (CNF) of the measurement distribution $f_{\boldsymbol{X}}(\boldsymbol{x};\boldsymbol{\theta})$ using the dataset $\mathcal{D}_T = \{\boldsymbol{x}_i, \boldsymbol{\theta}_i\}_{i=1}^{N_{\mathcal{D}}}$ of $N_{\mathcal{D}}$ measurement-parameter sample pairs.

Conditional Normalizing Flow (CNF)

- Train a conditional generator $\mathrm{G}(\boldsymbol{Z};\boldsymbol{\theta})$ that has a conditioning input $\boldsymbol{\theta}$ and random input $\boldsymbol{Z} \sim \mathcal{N}(0, \boldsymbol{I})$ to simulate the measurement process:

$$\Gamma(\boldsymbol{\theta}) \triangleq \mathrm{G}(\boldsymbol{Z};\boldsymbol{\theta}) \quad \text{approximately} \quad \sim f_{\boldsymbol{X}}(\boldsymbol{x};\boldsymbol{\theta}) \tag{3}$$

- The trained G (a neural network) is a deterministic function of $\boldsymbol{\theta}$ and $\boldsymbol{Z}$.
- $\boldsymbol{Z}$ is random $\implies$ the generative model is a random mapping from $\boldsymbol{\theta}$ to $\Gamma(\boldsymbol{\theta})$.
- G is invertible w.r.t. $\boldsymbol{Z}$, with inverse $\nu(\boldsymbol{\gamma}, \boldsymbol{\theta})$ (the normalizing flow).
- Training G: maximizes the likelihood of the data $\mathcal{D}_T$.
  Equivalent to minimizing (a sample estimate of) the KL divergence, for fixed $\boldsymbol{\theta}$, between the generated samples $\mathrm{G}(\boldsymbol{Z};\boldsymbol{\theta})$ and the true distribution $f_{\boldsymbol{X}}(\boldsymbol{x};\boldsymbol{\theta})$.

Approach

Stage 2: Use the learned conditional normalizing flow to approximate the Barankin bound. By relying on the invertibley of G and ν and the transformation of random variable:

$$f_{\Gamma}\left(\boldsymbol{\gamma};\boldsymbol{\theta}\right)=f_{\boldsymbol{Z}}\left(\nu\left(\boldsymbol{\gamma};\boldsymbol{\theta}\right)\right)\left|\det \boldsymbol{J}_{\nu}\left(\boldsymbol{\gamma};\boldsymbol{\theta}\right)\right| \tag{4}$$

Approximating the Barankin bound

We generate $\mathcal{D}_{\mathrm{G}}$ using G:

$$\mathcal{D}_{\mathrm{G}}=\left\{\boldsymbol{\gamma}_n=\mathrm{G}\left(\boldsymbol{z}_n;\boldsymbol{\theta}\right)\middle|\boldsymbol{\gamma}_n\in\tilde{\Upsilon},\boldsymbol{z}_n\sim\mathcal{N}(0,\boldsymbol{I})\right\}_{n=1}^{N_{\mathcal{D}_{\mathrm{G}}}}. \tag{5}$$

Using $\mathcal{D}_{\mathrm{G}}$ we compute the Barankin matrix using an empirical mean:

$$b\left(\boldsymbol{\psi}\right)\triangleq\mathbb{E}_{\boldsymbol{X}}\left[\eta^2\left(\boldsymbol{X};\boldsymbol{\theta},\Delta\right)\right]\approx\overline{b}\left(\Delta\right)\triangleq\underbrace{\frac{1}{\left|\mathcal{D}_{\mathrm{G}}\right|}\sum_{\boldsymbol{\gamma}\in\mathcal{D}_{\mathrm{G}}}}_{\text{Empirical Mean}}\underbrace{\tilde{\eta}^2\left(\boldsymbol{\gamma};\boldsymbol{\theta},\Delta\right)}_{\text{Approximated LR}}, \tag{6}$$

where $\tilde{\eta}\left(\boldsymbol{\gamma};\boldsymbol{\theta},\Delta\right)=\frac{f_{\Gamma}(\boldsymbol{\gamma};\boldsymbol{\theta}+\Delta)}{f_{\Gamma}(\boldsymbol{\gamma};\boldsymbol{\theta})}$ is likelihood ratio approximation.

Regions

To improve the stability of the learned Generative Barankin bound, we divide it into three regions:

- Asymptotic (Converge to the CRB)
- No Information
- Transition

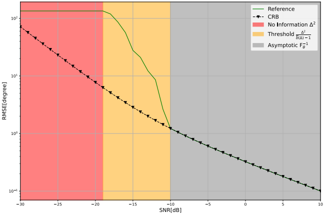

Figure: GBB Regions

Stable Generative Barankin Bound

$$\mathrm{BB}_s(\Delta) = \begin{cases} \mathrm{GCRB} & |\Delta| \leq \sqrt{\mathrm{GCRB}} \\ \Delta^2 & \overline{b}(\Delta) - 1 \leq 1 \quad \& \quad |\Delta| > \sqrt{\mathrm{GCRB}} \\ \frac{\Delta^2}{\overline{b}(\Delta)-1} & otherwise \end{cases}$$

where $\mathrm{GCRB}^a \approx \mathrm{CRB}$ is the Generative Cramér Rao bound.

[a] Habi, H. V., Messer, H., & Bresler, Y. (2023). Learning to bound: A generative Cramér-Rao bound. IEEE Transactions on Signal Processing. Also will be preseted in this ICASSP, Poster SPTM-P8.4, Fri, 19, 13:10 - 15:10

We demonstrate the Generative Barankin Bound on Direction Of Arrival (DOA) problems.

Direction-Of-Arrival Problem

Consider the case of a single source with uniform linear array (of size M):

$$\boldsymbol{X}_n = \boldsymbol{a}(\boldsymbol{\theta})\boldsymbol{S}_n + \boldsymbol{W}_n, \tag{7}$$

where,

- $[\boldsymbol{a}(\boldsymbol{\theta})]_m = \exp\left(\frac{2\pi}{\lambda} j x_m \sin(\boldsymbol{\theta})\right)$ is the steering vector.
- $\boldsymbol{S}_n \in \mathbb{C}$ is the random source signal at the snapshot n^{th}.
- $\boldsymbol{W}_n \in \mathbb{C}^M$ is an additive Gaussian noise with a covariance matrix $\boldsymbol{\Sigma_W}$

To improve sample complexity and convergence of our Normalizing Flow, we design a CNF (conditioned on the Direction of Arrival) specific for the DOA signal (DOAFlow). This is done by choosing a physics-informed approach that combines domain knowledge with learning.

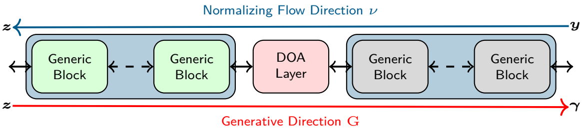

Figure: Architecture of the proposed DOA-Flow model.

- Generic Block: Includes standard CNF layers: Activation- Normalization, Invertible Linear Transformation (the so-called 1×1 convolution) and Coupling Layer.
- DOA layer: Incorporates the physical knowledge of the DOA problem into the DOA-Flow.

DOA Normalizing Flow Step

Let $\boldsymbol{z}_i$ and $\boldsymbol{z}_{i+1}$ be the input and output of the DOA layer, respectively. Then they are related by

$$\boldsymbol{z}_{i+1} = \boldsymbol{U}_{\boldsymbol{X}}\left(\boldsymbol{\theta}\right)^{\frac{1}{2}} \boldsymbol{z}_i \text{ and, } \boldsymbol{z}_i = \boldsymbol{U}_{\boldsymbol{X}}\left(\boldsymbol{\theta}\right)^{-\frac{1}{2}} \boldsymbol{z}_{i+1}, \tag{8}$$

where $\boldsymbol{U}_{\boldsymbol{X}}$ is the learned DOA layer covariance matrix

$$\boldsymbol{U}_{\boldsymbol{X}}\left(\boldsymbol{\theta}\right) = \boldsymbol{A}\left(\boldsymbol{\theta}\right)\boldsymbol{U}_S\boldsymbol{A}^H\left(\boldsymbol{\theta}\right) + \alpha\left(\mathrm{SNR}\right)\boldsymbol{U}_W \tag{9}$$

- $\alpha\left(\mathrm{SNR}\right) = 10^{-\mathrm{SNR}/10}$ is the SNR scale
- $\boldsymbol{U}_W$ and $\boldsymbol{U}_S$ are the learned noise and signal covariance matrices, respectively. To *ensure* that $\boldsymbol{U}_W$ and $\boldsymbol{U}_S$ are P.S.D. we represent them as LDU decomposition.
- $\boldsymbol{A}\left(\boldsymbol{\theta}\right)$ is the steering matrix (which is known based on the nominal sensor locations).

In the experiments, we investigate three scenarios:

- Gaussian signal and noise (reference scenario to validate the GBB).
- Sensor locations randomly perturbed.
- Constant amplitude signal (QAM4).

QAM4

$$S_n = \frac{1}{\sqrt{2}} \left(a_n + j \cdot b_n \right), \qquad (10)$$

where a_n and b_n are i.i.d. random variables taking the values ± 1 with equal probability 0.5.

Perturbed Sensor Locations

In this case the sensor locations: are

$$x_m = (m-1)\lambda/2 + U_m \qquad (11)$$

where $U_m \sim \mathcal{N}(0, \gamma^2)$ are i.i.d. This perturbation modifies the signal received by the array, but is unknown to the DOA estimator.

Test Point Selection

At the lowest SNR we set the farthest possible test point and sweep the SNR until $\mathrm{BB}_s < \mathrm{GCRB}$; then, for higher SNRs, we set the test point at $\psi = \theta$.

$$\psi = \begin{cases} \psi_0 & otherwise \\ \theta & \mathrm{BB}_s\left(\psi_0\right) < \mathrm{GCRB} \end{cases} \tag{12}$$

where $\psi_0 = \arg\max_{-\pi/2 \leq \psi \leq \pi/2} \left|\psi - \theta\right|$

In all experiment we use the following parameters: $M = 20$, number of snapshots $N = 5$, SNR range between -30 dB to 10 dB.

Gaussian

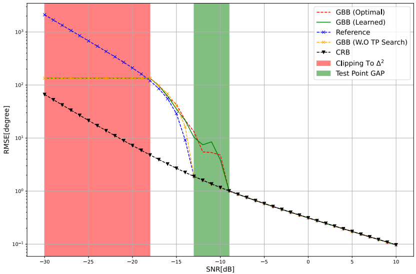

Figure: GBB: Gaussian Signal and Noise

non-Gaussian

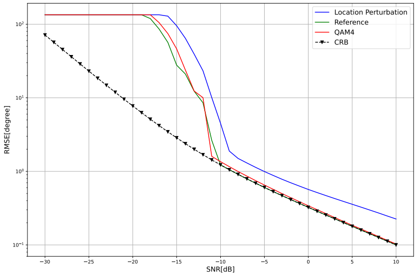

Figure: QAM4 Signal, Perturbed Sensor Locations and Reference Scenario (CRB+GBB)

Conclusions

- We have presented the Generative Barankin bound, the first learning-based bound on large estimation errors.
- We demonstrated its abilities and benefits on a DOA problem with a single source in three cases.
- We introduced *DOA-Flow*, a conditional normalizing flow for DOA signals.

Open Questions?

- How to select test points in the general case?
- Theoretical analysis of the convergence of the GBB to the BB ?