Bounds on Bearing, Symbol, and Channel Estimation under Model Misspecification

Akshay S. Bondre, Touseef Ali, and Christ D. Richmond

Center of Excellence Meeting

January 11th 2021, Monday, 12pm EDT

Signals, Information, Inference, & Learning (SIIL) Group

Arizona State University

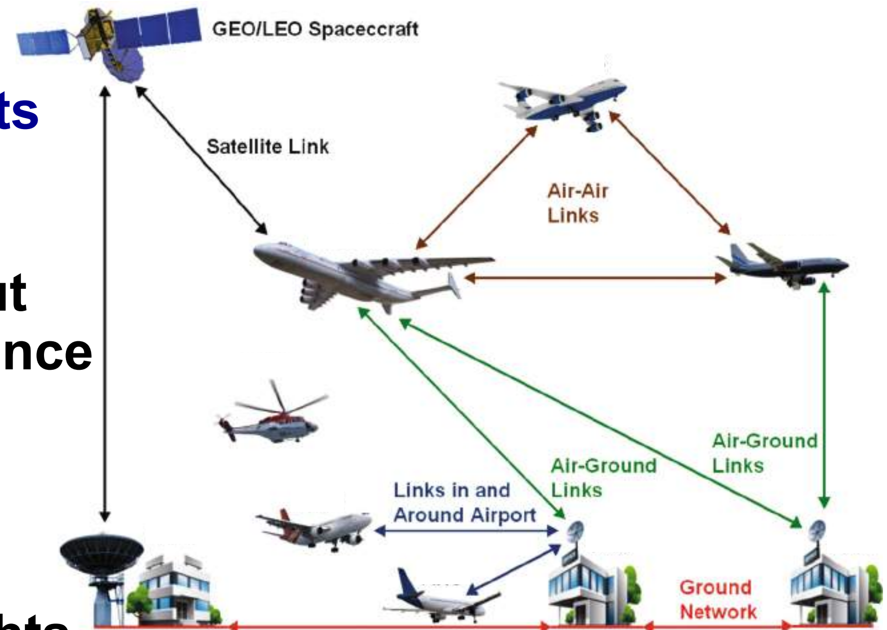
***website:** richmond.engineering.asu.edu

Outline

- Introduction
  - Remarks on channels, learning, and bounds
- Akshay's discussion
- Closing remarks

Remarks

- Interest is communication link performance in **dynamic environments**
- **Deep learning** is not model-based, but requires stationary data for convergence
  - Models may be leveraged, however, to improve convergence rates
- **Models** can also provide useful insights
- **Parameter estimation** plays key role in comm. link performance
  - **Maximum-likelihood decoding** often involves symbol, channel gain, bearing, Doppler, and delay estimation, etc.



- **Cramér-Rao bound** is useful for characterizing parameter estimation performance
  - But comm. links present **some challenges**
    - Symbols are discrete
    - Model is assumed known

Cramér-Rao Bound: Scalar Parameters

- **Recall scalar covariance inequality:**

$$1 \geq \frac{E^2\{\zeta\eta\}}{E\{\zeta^2\}E\{\eta^2\}} \quad \Rightarrow \quad E\{\zeta^2\} \geq \frac{E^2\{\zeta\eta\}}{E\{\eta^2\}}$$

- **Consider parameter estimate*:** $\widehat{\theta} = \widehat{\theta}(\mathbf{x})$, $E\{\widehat{\theta}\} = \theta$, $\mathbf{x} \sim p(\mathbf{x}|\theta_T)$

- **Choose random variables (r.v.) as:** $\zeta = \widehat{\theta} - \theta$, $\eta = \frac{\partial \ln p}{\partial \theta}$ ***Score Function***

- **Inequality can lower bound estimator mean squared error (MSE):**

(Translation)^2

$$E\{(\widehat{\theta} - \theta)^2\} \geq \frac{\boxed{E^2\left\{(\widehat{\theta} - \theta)\frac{\partial \ln p}{\partial \theta}\right\}}}{E\left\{\left(\frac{\partial \ln p}{\partial \theta}\right)^2\right\}} = \frac{1}{E\left\{\left(\frac{\partial \ln p}{\partial \theta}\right)^2\right\}\Bigg|_{\theta=\theta_T}} = \mathrm{CRB}(\theta_T)$$

Cramér-Rao Bound (CRB)

- **CRB given by inverse Fisher Information:** $\mathrm{FIM}(\theta) = E\left\{\left(\frac{\partial \ln p}{\partial \theta}\right)^2\right\}$

- **CRB applies to all unbiased estimators, i.e. such that** $E\{\widehat{\theta}\} = \theta$

***All expectations with respect to distribution of data x.**

Cramér-Rao Bound: Vector Parameters

- **Let** $\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_M]^T$**,** $\widehat{\boldsymbol{\theta}}(\mathbf{x}) = [\widehat{\theta}_1(\mathbf{x}), \widehat{\theta}_2(\mathbf{x}), \ldots, \widehat{\theta}_M(\mathbf{x})]^T$**;** $E\{\widehat{\boldsymbol{\theta}}(\mathbf{x})\} = \boldsymbol{\theta}$
- **Multivariate covariance inequality:**

$$E\{\boldsymbol{\zeta}\boldsymbol{\zeta}^T\} \geq E\{\boldsymbol{\zeta}\boldsymbol{\eta}^T\}E^{-1}\{\boldsymbol{\eta}\boldsymbol{\eta}^T\}E\{\boldsymbol{\eta}\boldsymbol{\zeta}^T\}$$

- **Variable choice for** $\boldsymbol{\zeta}$ **and** $\boldsymbol{\eta}$ **lower bounds estimator MSE matrix:**

$$\boldsymbol{\zeta} = \widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}, \quad \boldsymbol{\eta} = \frac{\partial \ln p(\mathbf{x}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \Rightarrow \begin{array}{c} E\{\boldsymbol{\zeta}\boldsymbol{\zeta}^T\} = \mathrm{Cov}(\widehat{\boldsymbol{\theta}}) = \\ E\left\{(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T\right\} \geq \mathbf{CRB}(\boldsymbol{\theta}) \end{array}$$

- **Fisher information matrix (FIM) and translation matrix:**

$$\mathbf{J}(\boldsymbol{\theta}) \triangleq E\left\{\frac{\partial \ln p(\mathbf{x}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial \ln p(\mathbf{x}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}^T\right\}, \quad \boldsymbol{\Xi}(\boldsymbol{\theta}) \triangleq E\left\{\frac{\partial \ln p(\mathbf{x}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T\right\} = \mathbf{I}_M$$

$E\{\boldsymbol{\eta}\boldsymbol{\eta}^T\}$ $E\{\boldsymbol{\eta}\boldsymbol{\zeta}^T\}$

- **CRB lower bound for unbiased estimators given by:**

$$\boxed{E\left\{(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T\right\} \geq \mathbf{J}^{-1}(\boldsymbol{\theta})} \Rightarrow E\{[\widehat{\theta}_i(\mathbf{x}) - \theta_i]^2\} \geq \left[\mathbf{J}^{-1}(\boldsymbol{\theta})\right]_{i,i}$$

True parameter vector** θ **assumed real and deterministic but unknown. Data** x **may be complex or real.

Cramér-Rao Bound: FIM Slepian Formula

- **Let parameters be given by** $\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_M]^T$

- **Assume data is complex Gaussian, i.e.** $\tilde{\mathrm{x}} \sim \mathcal{CN}(\boldsymbol{\mu}(\boldsymbol{\theta}), \tilde{\mathbf{R}}) = p(\tilde{\mathrm{x}}|\boldsymbol{\theta})$
  - **Only mean is parameterized by** θ

- **Slepian formula for FIM, denoted** $\mathbf{J}(\boldsymbol{\theta})$**, is given by**

$$\mathbf{J}(\boldsymbol{\theta}) = \frac{\partial \boldsymbol{\mu}^*}{\partial \boldsymbol{\theta}^*} \tilde{\mathbf{R}}^{-1} \left(\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\theta}} \right)^T + \frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\theta}^*} \left(\tilde{\mathbf{R}}^{-1} \right)^* \left(\frac{\partial \boldsymbol{\mu}^*}{\partial \boldsymbol{\theta}} \right)^T$$

- **Cramér-Rao bound:**

$$\boxed{\mathbb{E}\left[\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta} \right) \left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta} \right)^H \right] \geq \mathbf{J}^{-1}(\boldsymbol{\theta})}$$

Constrained Cramér-Rao Bound

1. **Given m constraints on $\boldsymbol{\theta}$:** $\mathbf{f}(\boldsymbol{\theta}) = \left[f_1(\boldsymbol{\theta}), \ldots, f_m(\boldsymbol{\theta})\right]^T = \mathbf{0}_{m\times 1}$

2. **Gradient matrix $\mathbf{F}(\boldsymbol{\theta})$:** $\mathbf{F}(\boldsymbol{\theta}) = \left[\frac{\partial \mathbf{f}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^*}\right]^T$

3. **There exists a unitary matrix $\mathbf{U}_{M\times(M-m)}$ such that**

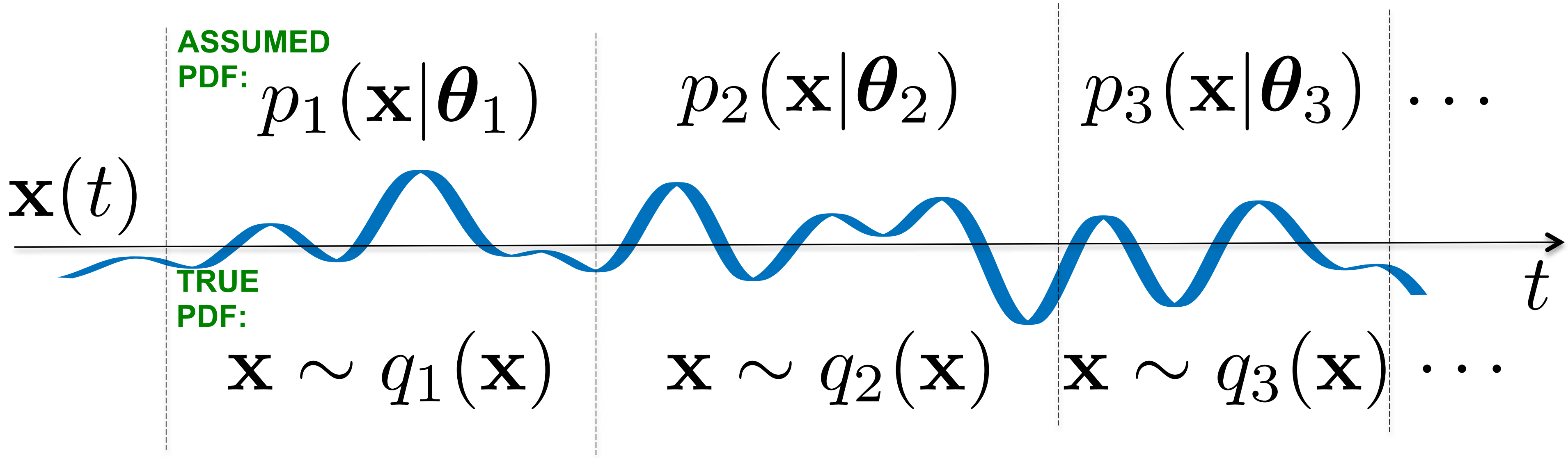
$$\mathbf{F}(\boldsymbol{\theta})\mathbf{U} = \mathbf{0}_{m\times(M-m)}$$

4. **CCRB**

$$\boxed{\mathbb{E}\left[\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right)\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right)^H\right] \geq \mathbf{U}(\mathbf{U}^H\mathbf{J}(\boldsymbol{\theta})\mathbf{U})^{-1}\mathbf{U}^H}$$

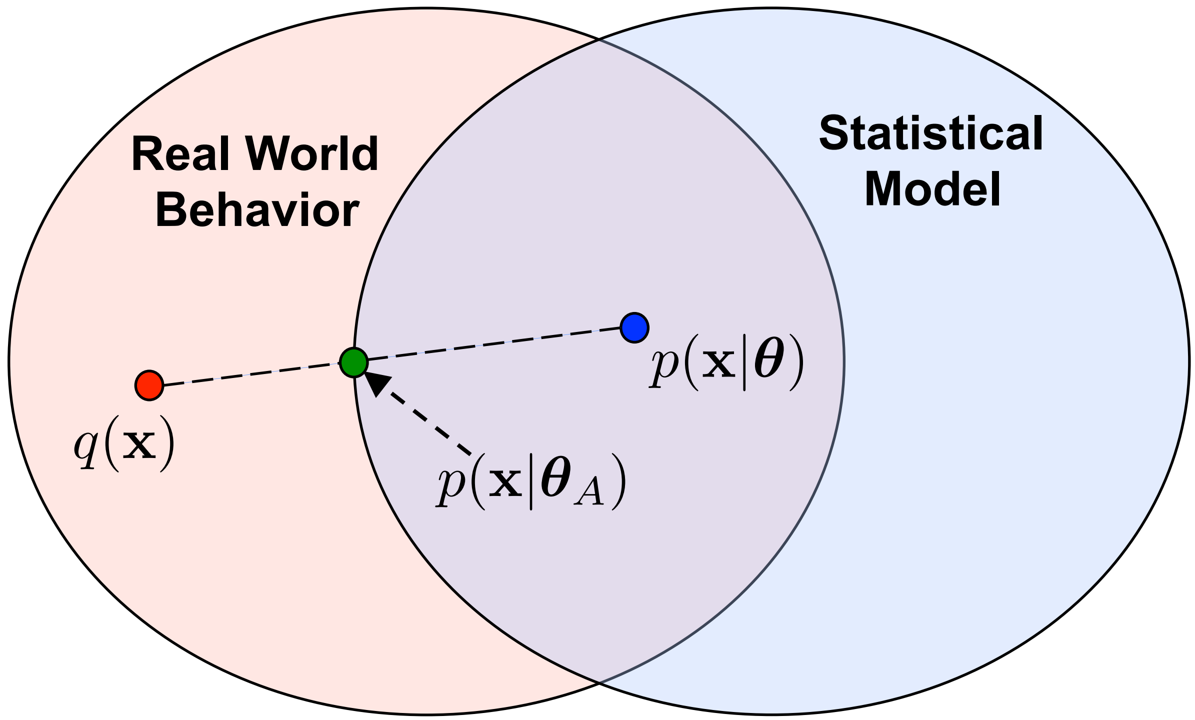
Will use CCRB to address any discrete parameters

Remarks on Dynamic Comm. Channels



- **Possible initial framework for modeling dynamic channels**
- **Simulated data can feed deep learning algorithms**
- **Models can provide metrics by which to measure performance**
- **Models may help resolve/improve deep learning performance**

Model Misspecification*



- **Consider $\mathbf{x} \sim q(\mathbf{x}) \neq p(\mathbf{x}|\boldsymbol{\theta})$ for all θ is allowed**
  - **Distance between $q(\mathbf{x})$ and $p(\mathbf{x}|\boldsymbol{\theta})$ may be bounded away from zero**
- **What can be said about limits of parameter estimation?**
  - **How well can we expect to do?**

S. Fortunati, F. Gini, M. S. Greco, and C. D. Richmond, IEEE Signal Processing Magazine, vol. 34, no. 6, pp. 142–157, Nov 2017.

Ira A. Fulton Schools of Engineering
Arizona State University

Misspecified Cramér-Rao Bound (MCRB)

- **Consider complex data $\tilde{\mathrm{x}}$ with distributions:**

  True: $\tilde{\mathrm{x}} \sim q(\tilde{\mathrm{x}}) = \mathcal{CN}(\mathrm{d}, \tilde{\mathrm{R}})$, **Assumed:** $\tilde{\mathrm{x}} \sim p(\tilde{\mathrm{x}}|\boldsymbol{\theta}) = \mathcal{CN}(\boldsymbol{\mu}(\boldsymbol{\theta}), \tilde{\mathrm{R}})$

  – ***Mean is parameterized*** **and $\boldsymbol{\mu}(\boldsymbol{\theta}) \neq \mathrm{d}$ for all $\boldsymbol{\theta}$ is allowed***

- **Define $\delta\boldsymbol{\mu}(\boldsymbol{\theta}) \triangleq \mathrm{d} - \boldsymbol{\mu}(\boldsymbol{\theta})$, and matrices:**

$$\mathbf{J}(\boldsymbol{\theta}) \triangleq \left(\frac{\partial \boldsymbol{\mu}^*(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^*}\right) \tilde{\mathbf{R}}^{-1} \left(\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right)^T + \left(\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^*}\right) (\tilde{\mathbf{R}}^{-1})^* \left(\frac{\partial \boldsymbol{\mu}^*(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right)^T$$

$$[\mathbf{C}(\boldsymbol{\theta})]_{i,k} \triangleq -\left[\frac{\partial \boldsymbol{\mu}^*(\boldsymbol{\theta})}{\partial \theta_i^*}\right]^T \tilde{\mathbf{R}}^{-1} \frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_k} - \left[\frac{\partial \boldsymbol{\mu}^*(\boldsymbol{\theta})}{\partial \theta_k}\right]^T \tilde{\mathbf{R}}^{-1} \frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_i^*}$$
$$+ \left[\frac{\partial^2 \boldsymbol{\mu}^*(\boldsymbol{\theta})}{\partial \theta_i^* \partial \theta_k}\right]^T \tilde{\mathbf{R}}^{-1} \delta\boldsymbol{\mu}(\boldsymbol{\theta}) + \delta\boldsymbol{\mu}^H(\boldsymbol{\theta}) \tilde{\mathbf{R}}^{-1} \frac{\partial^2 \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_i^* \partial \theta_k}$$

- **Let $\boldsymbol{\theta}_A \triangleq \arg\min_{\boldsymbol{\theta}} D(q_{\tilde{\mathbf{x}}} \,||\, p_{\tilde{\mathbf{x}}|\boldsymbol{\theta}})$ and $E_q\{\widehat{\boldsymbol{\theta}}\} = \boldsymbol{\theta}_A$. The MCRB is given by**:**

$$\boxed{E_q\left\{(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_A)(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_A)^H\right\} \geq \mathbf{C}^{-1}(\boldsymbol{\theta}_A)\mathbf{J}(\boldsymbol{\theta}_A)\mathbf{C}^{-1}(\boldsymbol{\theta}_A)}$$

***Parameters θ can be real and/or complex. Wirtinger calculus used.**

*****Richmond and Horowitz, IEEE T-SP, Vol. 63, No. 9, May 2015***

Classes of Cramér-Rao Lower Bounds

Cramér-Rao Bound (CRB)
(Continuous parameters)

Q.H. Vuong,, Division Humanities Soc. Sci., Caltech, Oct 1986.

Misspecified CRB (MCRB)
(CRB under model misspecification)

C.D. Richmond and L.L. Horowitz, IEEE Trans. Sig. Proc., May 2015.

S. Fortunati, F. Gini, M. S. Greco, and C. D. Richmond, *IEEE Signal Processing Magazine*, vol. 34, no. 6, pp, Nov. 2017.

P. Stoica and B.C. Ng, IEEE Signal Processing Letters, July 1998.

Constrained CRB (CCRB)
(Side information can be used for discrete comm. signals)

A.K. Jagannatham and B. D. Rao, *IEEE Signal Processing Letters*, Nov 2004.

Sadler, Kozick, and Moore, IEEE Trans. On Sig. Proc., vol. 49, no. 4, 2001.

Constrained Misspecified CRB (CMCRB)
(Comm. signals under model misspecification)

S. Fortunati, F. Gini, M.S. Greco, IEEE Signal Processing Letters, vol.23, no. 5, 2016.

Outline

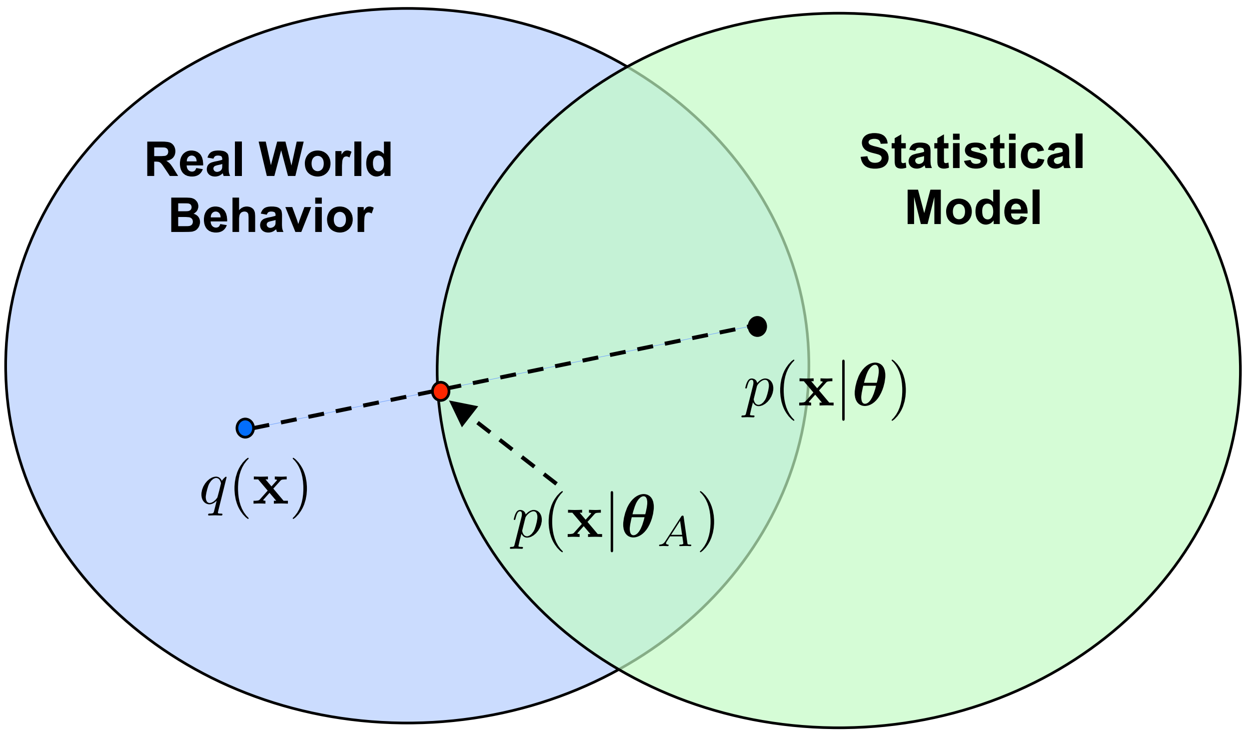
- Introduction
- **Akshay's discussion**
- Closing remarks

Akshay Outline

- **Introduction**
- Communications Data Model
- Constrained Misspecified Cramer-Rao Bound
- Numerical Results
- Summary

Introduction

- **Model Misspecification**



  - **Bounds on accuracy in estimating parameter θ.**

Introduction

- <u>GOAL</u>: Develop bounds on parameter estimation performance in a communication link under model misspecification.

- Model misspecification: | Assumed model $\neq$ True Model |

- Cramer-Rao bound $\Longrightarrow$ CONTINUOUS PARAMETER.

- <u>PROBLEM</u>:

  Communication signals $\Longrightarrow$ Discrete symbols

  (Finite parameter space)

Ira A. Fulton Schools of Engineering
Arizona State University

Cramer-Rao Lower Bound

Cramer-Rao Bound (CRB)
(Continuous parameters)

Misspecified CRB (MCRB)
(CRB under model misspecification)

Q.H. Vuong,, Division Humanities Soc. Sci., Caltech, Oct 1986.
C.D. Richmond and L.L. Horowitz, IEEE Trans. Sig. Proc., May 2015.

Constrained CRB (CCRB)
(Side information can be used for discrete comm. signals)

P. Stoica and B.C. Ng, IEEE Signal Processing Letters, July 1998.
Sadler, Kozick, and Moore, IEEE Trans. On Sig. Proc., vol. 49, no. 4, 2001.

Constrained Misspecified CRB (CMCRB)
(Comm. signals under model misspecification)

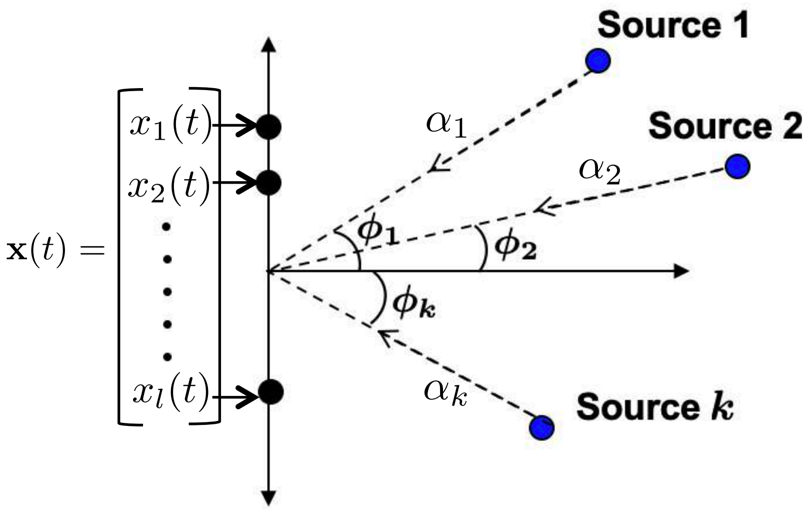
S. Fortunati, F. Gini, M.S. Greco, IEEE Signal Processing Letters, vol.23, no. 5, 2016.

Ira A. Fulton Schools of Engineering
Arizona State University

Akshay Outline

- Introduction
- **Communications Data Model**
- Constrained Misspecified Cramer-Rao Bound
- Numerical Results
- Summary

Flat-fading Data Model



- l sensors and k sources.
- ϕ_i: Angles of arrival, $\boldsymbol{\phi} = \left[\phi_1, \ldots, \phi_k\right]^T$
- α_i: Channel gains, $\boldsymbol{\alpha} = \left[\alpha_1, \ldots, \alpha_k\right]^T$
- $\mathbf{a}(\phi_i)$: Steering vectors
- $s_i(t)$: Comm. signal from source i at time t, $t = 1, \ldots, N.$

$$\mathbf{s}(t) = \left[s_1(t), \ldots, s_k(t)\right]^T$$

- **Flat-fading model**: Channel gains $\implies$ constant over time, ($\alpha_i(t) = \alpha_i$)

Received signal $\mathbf{x}(t)$:

$$\boxed{\mathbf{x}(t) = \sum_{i=1}^{k} \mathbf{a}(\phi_i)\alpha_i s_i(t) + \mathbf{n}(t) = \mathbf{A}(\boldsymbol{\phi})\boldsymbol{\Delta}(\boldsymbol{\alpha})\mathbf{s}(t) + \underbrace{\mathbf{n}(t)}_{\mathcal{CN}(\mathbf{0}, \mathbf{R})}}$$

Flat-fading Data Model

- $$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_l(t) \end{bmatrix}_{l\times 1} = \overbrace{\begin{bmatrix} | & & | \\ \mathbf{a}(\phi_1) & \cdots & \mathbf{a}(\phi_k) \\ | & & | \end{bmatrix}_{l\times k}}^{\mathbf{A}(\boldsymbol{\phi})} \cdot \overbrace{\begin{bmatrix} \alpha_1 & 0 & \cdots & 0 \\ 0 & \alpha_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \alpha_k \end{bmatrix}_{k\times k}}^{\boldsymbol{\Delta}(\boldsymbol{\alpha})} \cdot \begin{bmatrix} s_1(t) \\ s_2(t) \\ \vdots \\ s_k(t) \end{bmatrix}_{k\times 1} + \begin{bmatrix} n_1(t) \\ n_2(t) \\ \vdots \\ n_l(t) \end{bmatrix}_{l\times 1}$$
$$= \mathbf{A}(\boldsymbol{\phi})\boldsymbol{\Delta}(\boldsymbol{\alpha})\mathbf{s}(t) + \boxed{\mathbf{n}(t)} \longleftarrow \mathcal{CN}(\mathbf{0}, \mathbf{R})$$

- N **such snapshots are stacked to form the data vector $\tilde{\mathbf{x}}$.**

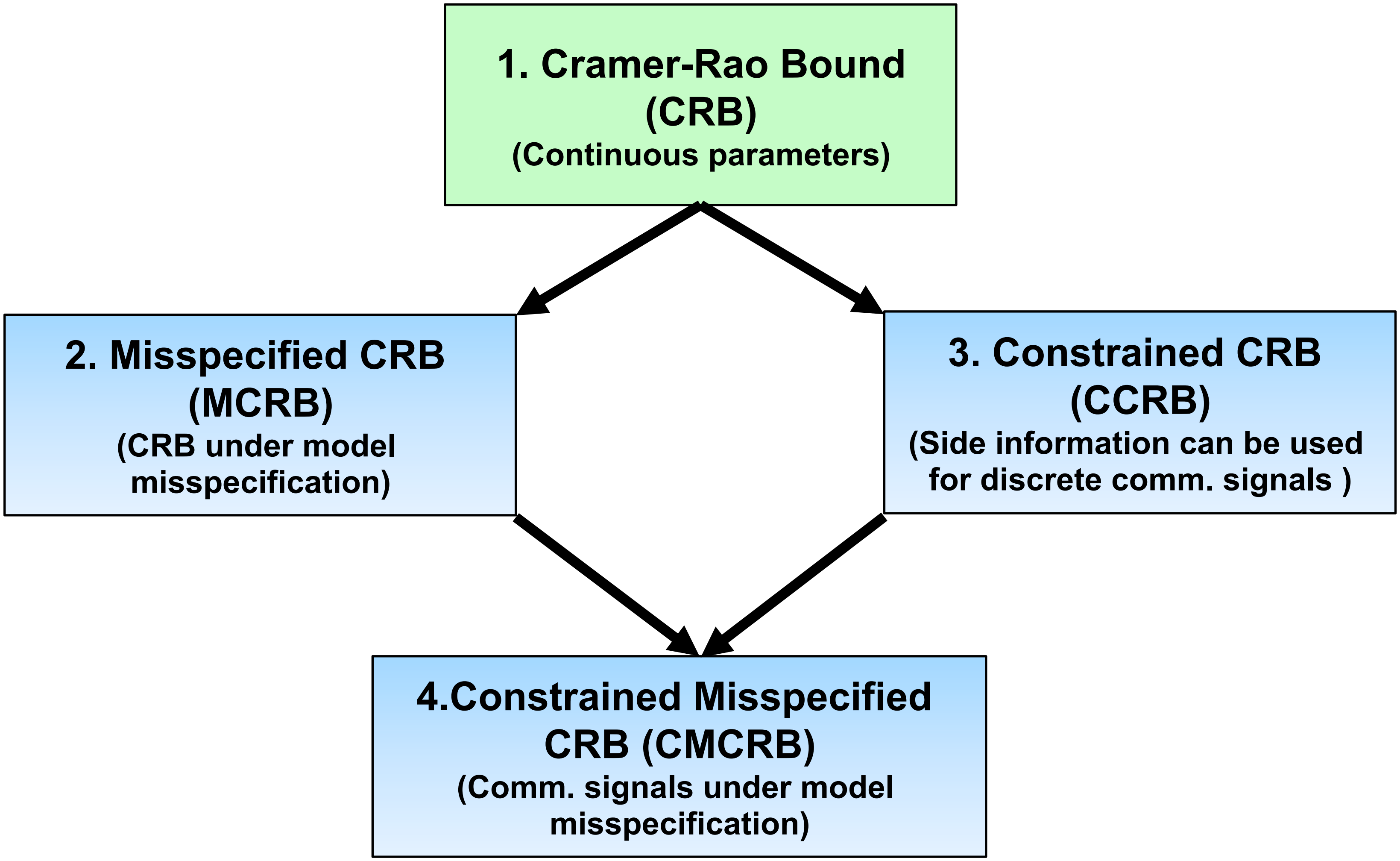
- $$\tilde{\mathbf{x}} = \begin{bmatrix} \mathbf{x}(1) \\ \vdots \\ \mathbf{x}(N) \end{bmatrix} = \begin{bmatrix} \mathbf{A}(\boldsymbol{\phi})\boldsymbol{\Delta}(\boldsymbol{\alpha})\mathbf{s}(1) \\ \vdots \\ \mathbf{A}(\boldsymbol{\phi})\boldsymbol{\Delta}(\boldsymbol{\alpha})\mathbf{s}(N) \end{bmatrix} + \begin{bmatrix} \mathbf{n}(1) \\ \vdots \\ \mathbf{n}(N) \end{bmatrix} = \boldsymbol{\mu} + \boxed{\tilde{\mathbf{n}}} \longleftarrow \mathcal{CN}(\mathbf{0}, \underbrace{\mathbf{I}_N \otimes \mathbf{R}}_{=\tilde{\mathbf{R}}})$$

$$\boxed{\tilde{\mathbf{x}} \sim \mathcal{CN}(\boldsymbol{\mu}, \tilde{\mathbf{R}})}$$

Akshay Outline

- Introduction
- Communications Data Model
- **Constrained Misspecified Cramer-Rao Bound**
- Numerical Results
- Summary

Cramer-Rao Lower Bound



S. Fortunati, F. Gini, M.S. Greco, "The constrained misspecified Cramer-Rao bound", IEEE Signal Processing Letters, vol.23, no. 5, 2016.

Unconstrained Cramer-Rao Bound

- **Parameter vector** $\boldsymbol{\theta} = \left[\theta_1, \ldots, \theta_M\right]^T = \left[\tilde{\mathbf{s}}; \tilde{\mathbf{s}}^*; \boldsymbol{\alpha}; \boldsymbol{\alpha}^*; \boldsymbol{\phi}\right]^{\boldsymbol{T}}$.

  $\tilde{\mathbf{s}} = \left[\mathbf{s}(1); \mathbf{s}(2); \ldots; \mathbf{s}(N)\right]^T$.

$$\boldsymbol{\alpha} = \left[\alpha_1, \ldots, \alpha_k\right]^T$$
$$\boldsymbol{\phi} = \left[\phi_1, \ldots, \phi_k\right]^T$$
$$\mathbf{s}(t) = \left[s_1(t), \ldots, s_k(t)\right]^T$$

- $\tilde{\mathbf{x}} \sim \mathcal{CN}(\boldsymbol{\mu}(\boldsymbol{\theta}), \tilde{\mathbf{R}}) = p_{\tilde{\mathbf{x}}|\boldsymbol{\theta}}$

  $\Longrightarrow$ **Only the mean is characterized by** $\boldsymbol{\theta}$.

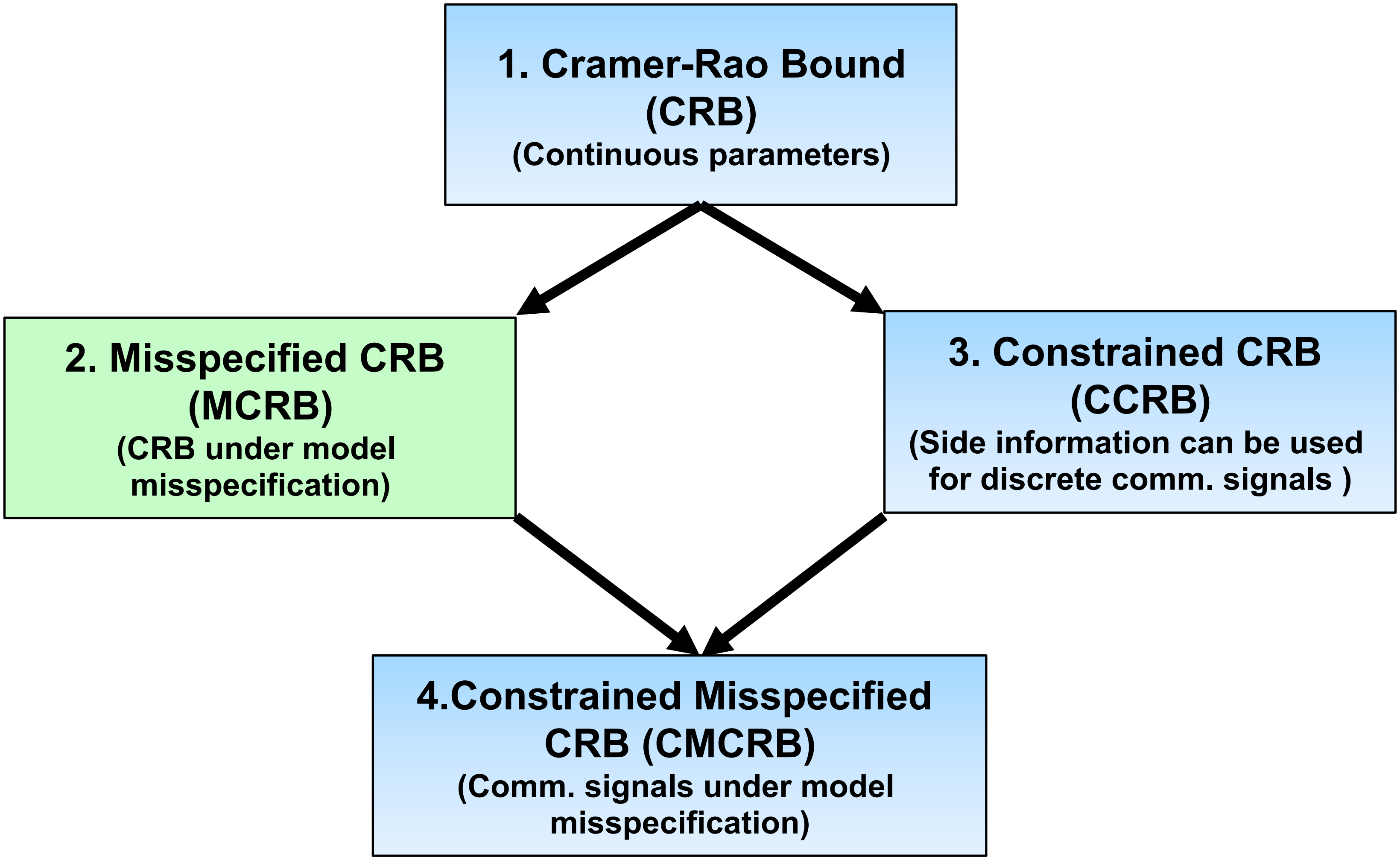
- **Unconstrained Fisher Information Matrix (FIM)** $\mathbf{J}(\boldsymbol{\theta})$

$$\mathbf{J}(\boldsymbol{\theta}) = \frac{\partial \boldsymbol{\mu}^*}{\partial \boldsymbol{\theta}^*} \tilde{\mathbf{R}}^{-1} \left(\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\theta}}\right)^T + \frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\theta}^*} \left(\tilde{\mathbf{R}}^{-1}\right)^* \left(\frac{\partial \boldsymbol{\mu}^*}{\partial \boldsymbol{\theta}}\right)^T$$

- **Cramer-Rao Bound:**

$$\boxed{\mathbb{E}\left[\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right)\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right)^H\right] \geq \mathbf{J}^{-1}(\boldsymbol{\theta})}$$

Cramer-Rao Lower Bound



Misspecified Cramer-Rao Bound

- **Assumed flat-fading model:** $\tilde{\mathbf{x}} \sim \mathcal{CN}(\boldsymbol{\mu}(\boldsymbol{\theta}), \tilde{\mathbf{R}}) = p_{\tilde{\mathbf{x}}|\boldsymbol{\theta}}$
- **True data model:** $\tilde{\mathbf{x}} \sim \mathcal{CN}(\mathbf{d}, \tilde{\mathbf{R}}) = q_{\tilde{\mathbf{x}}}$

$\tilde{\mathbf{R}}$ **is assumed to be known and correct.**

$$\boxed{\boldsymbol{\theta}_A = \arg\min_{\boldsymbol{\theta}} D(q_{\tilde{\mathbf{x}}} \,||\, p_{\tilde{\mathbf{x}}|\boldsymbol{\theta}})}$$

- **Complexified Misspecified Fisher Information Matrix (MFIM)**

$$\mathbf{J}(\boldsymbol{\theta}) = \frac{\partial \boldsymbol{\mu}^*}{\partial \boldsymbol{\theta}^*} \tilde{\mathbf{R}}^{-1} \left(\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\theta}} \right)^T + \frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\theta}^*} \left(\tilde{\mathbf{R}}^{-1} \right)^* \left(\frac{\partial \boldsymbol{\mu}^*}{\partial \boldsymbol{\theta}} \right)^T$$

for $\tilde{\mathrm{R}}$ assumed to be known and correct (same as unconstrained FIM for this case).

Misspecified Cramer-Rao Bound

- **Complexified Average Hessian $\mathrm{C}(\boldsymbol{\theta})$**

$$[\mathbf{C}(\boldsymbol{\theta})]_{i,j} \triangleq -\left[\frac{\partial \boldsymbol{\mu}^*(\boldsymbol{\theta})}{\partial \theta_i^*}\right]^T \tilde{\mathbf{R}}^{-\mathbf{1}} \frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_j} - \left[\frac{\partial \boldsymbol{\mu}^*(\boldsymbol{\theta})}{\partial \theta_j}\right]^T \tilde{\mathbf{R}}^{-\mathbf{1}} \frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_i^*} + \left[\frac{\partial^2 \boldsymbol{\mu}^*(\boldsymbol{\theta})}{\partial \theta_i^* \partial \theta_j}\right]^T \tilde{\mathbf{R}}^{-\mathbf{1}} \delta\boldsymbol{\mu}(\boldsymbol{\theta}) + \delta\boldsymbol{\mu}^H(\boldsymbol{\theta}) \tilde{\mathbf{R}}^{-\mathbf{1}} \frac{\partial^2 \boldsymbol{\mu}^*(\boldsymbol{\theta})}{\partial \theta_i^* \partial \theta_j}$$

- **Misspecified CRB**

$$\boxed{\boldsymbol{\theta}_A = \arg\min_{\boldsymbol{\theta}} D(q_{\tilde{\mathbf{x}}} \,||\, p_{\tilde{\mathbf{x}}|\boldsymbol{\theta}})}$$

$$\boxed{\mathbb{E}\left[\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_A\right)\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_A\right)^H\right] \geq \mathbf{C}^{-1}(\boldsymbol{\theta}_{\boldsymbol{A}})\mathbf{J}(\boldsymbol{\theta}_A)\mathbf{C}^{-1}(\boldsymbol{\theta}_{\boldsymbol{A}})}$$

Cramer-Rao Lower Bound

1. Cramer-Rao Bound (CRB)
(Continuous parameters)

2. Misspecified CRB (MCRB)
(CRB under model misspecification)

3. Constrained CRB (CCRB)
(Side information can be used for discrete comm. signals)

4. Constrained Misspecified CRB (CMCRB)
(Comm. signals under model misspecification)

S. Fortunati, F. Gini, M.S. Greco, "The constrained misspecified Cramer-Rao bound", IEEE Signal Processing Letters, vol.23, no. 5, 2016.

Constrained Cramer-Rao Bound

1. **Given** m **constraints on** $\boldsymbol{\theta}$**:** $\mathbf{f}(\boldsymbol{\theta}) = \left[f_1(\boldsymbol{\theta}), \ldots, f_m(\boldsymbol{\theta})\right]^T = \mathbf{0}_{m\times 1}$

2. **Gradient matrix** $\mathbf{F}(\boldsymbol{\theta})$**:** $\mathbf{F}(\boldsymbol{\theta}) = \left[\frac{\partial \mathbf{f}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^*}\right]^T$

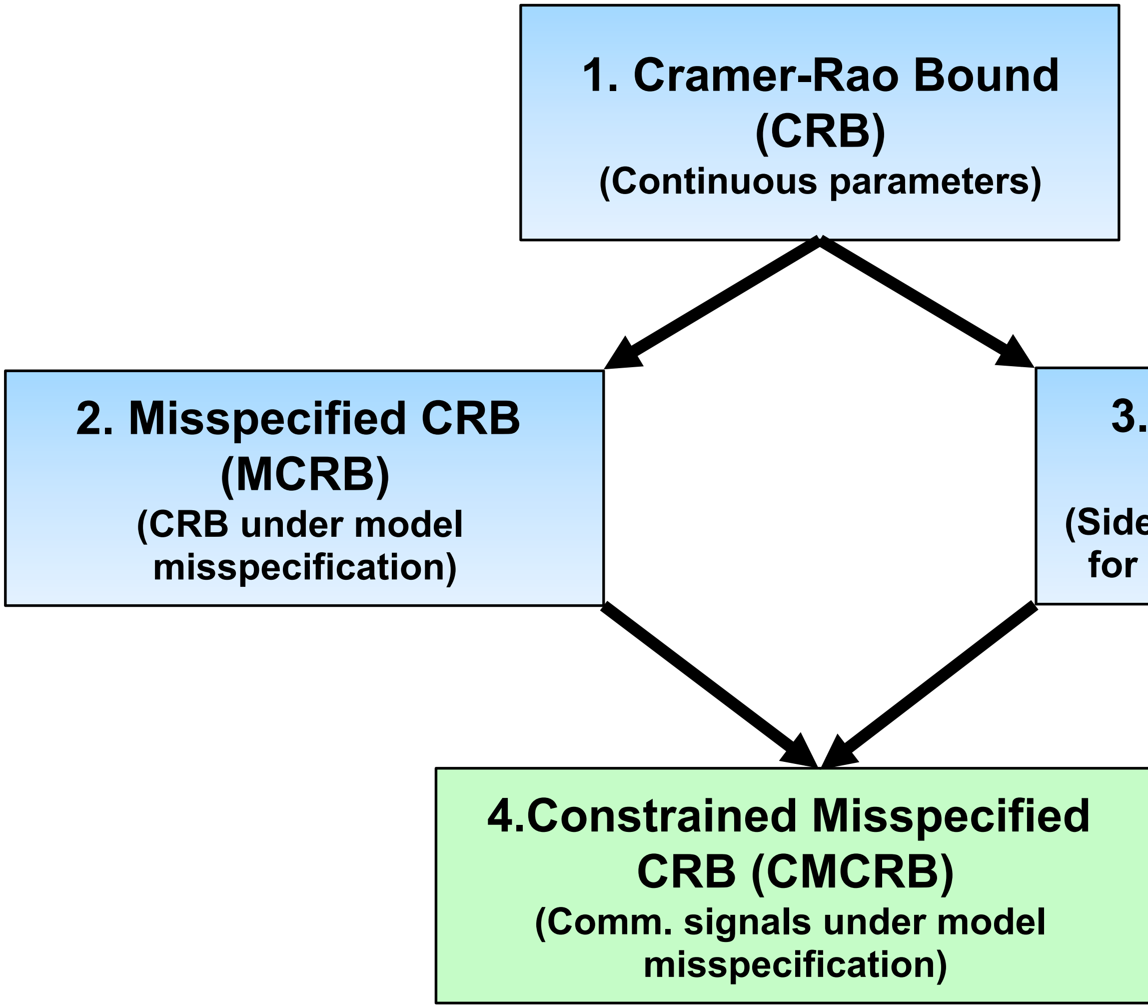
3. **There exists a unitary matrix** $\mathbf{U}$ **such that**

$$\mathbf{F}(\boldsymbol{\theta})\mathbf{U} = \mathbf{0}_{m\times(2Nk+3k-m)}$$

4. **CCRB**

$$\mathbb{E}\left[\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right)\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right)^H\right] \geq \mathbf{U}(\mathbf{U}^H\mathbf{J}(\boldsymbol{\theta})\mathbf{U})^{-1}\mathbf{U}^H$$

Cramer-Rao Lower Bound



S. Fortunati, F. Gini, M.S. Greco, "The constrained misspecified Cramer-Rao bound", IEEE Signal Processing Letters, vol.23, no. 5, 2016.

Constrained Misspecified Cramer-Rao Bound

- **Misspecified CRB (MCRB):**

$$\mathbb{E}\left[\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_A\right)\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_A\right)^H\right]\geq \mathbf{C}^{-1}(\boldsymbol{\theta}_{\boldsymbol{A}})\mathbf{J}(\boldsymbol{\theta}_A)\mathbf{C}^{-1}(\boldsymbol{\theta}_{\boldsymbol{A}})$$

- **Constrained CRB (CCRB):**

$$\mathbb{E}\left[\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)^H\right]\geq \mathbf{U}(\mathbf{U}^H\mathbf{J}(\boldsymbol{\theta})\mathbf{U})^{-1}\mathbf{U}^H$$

- **Constrained Misspecified CRB (CMCRB):**

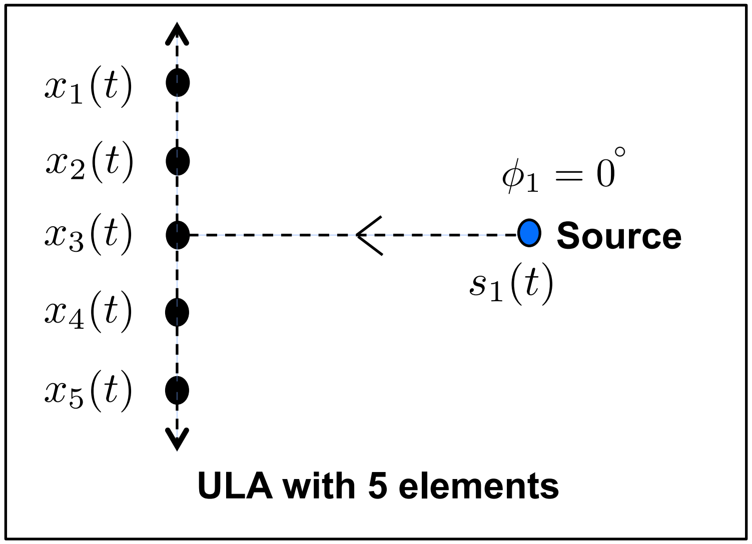
$$\boxed{\begin{aligned}\mathbb{E}\left[\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_A\right)\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_A\right)^H\right] &\geq \mathbf{U}(\mathbf{U}^H\mathbf{C}(\boldsymbol{\theta}_A)\mathbf{U})^{-1}\mathbf{U}^H \\ &\times \mathbf{J}(\boldsymbol{\theta}_A)\mathbf{U}(\mathbf{U}^H\mathbf{C}(\boldsymbol{\theta}_A)\mathbf{U})^{-1}\mathbf{U}^H\end{aligned}}$$

Akshay Outline

- Introduction
- Communications Data Model
- Constrained Misspecified Cramer-Rao Bound
- **Numerical Results**
- Summary

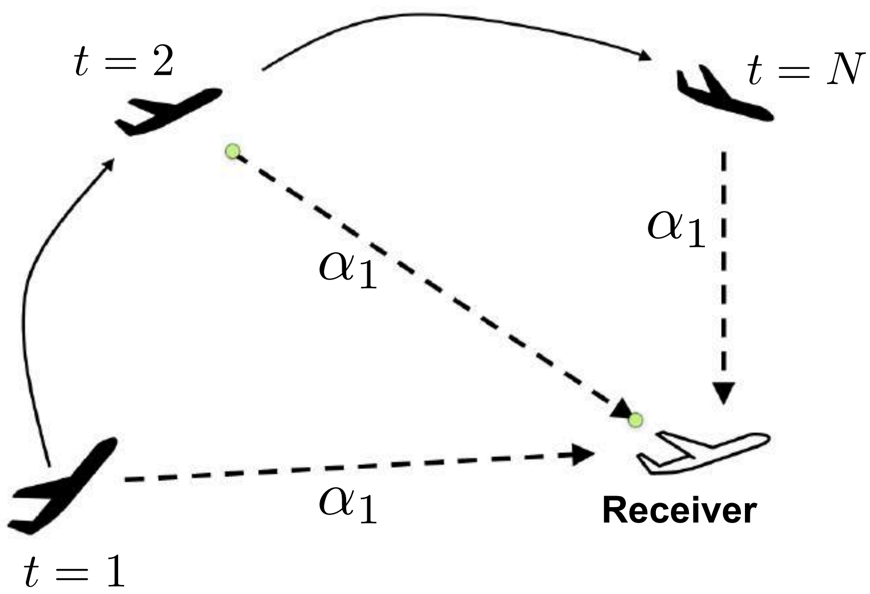
Simulation Setup

- **Single source** ⇒ $k = 1$
- **Angle of arrival** ⇒ $\phi_1 = 0^\circ$
- **Uniform Linear Array with 5 sensors** ⇒ $l = 5$
- **BPSK modulation with equal probability** ⇒ $s_1(t) = \pm 1$, $t = 1, \ldots, N = 100$
- **White Gaussian Noise** ⇒ $\tilde{\mathbf{R}} = \sigma^2 \mathbf{I}$

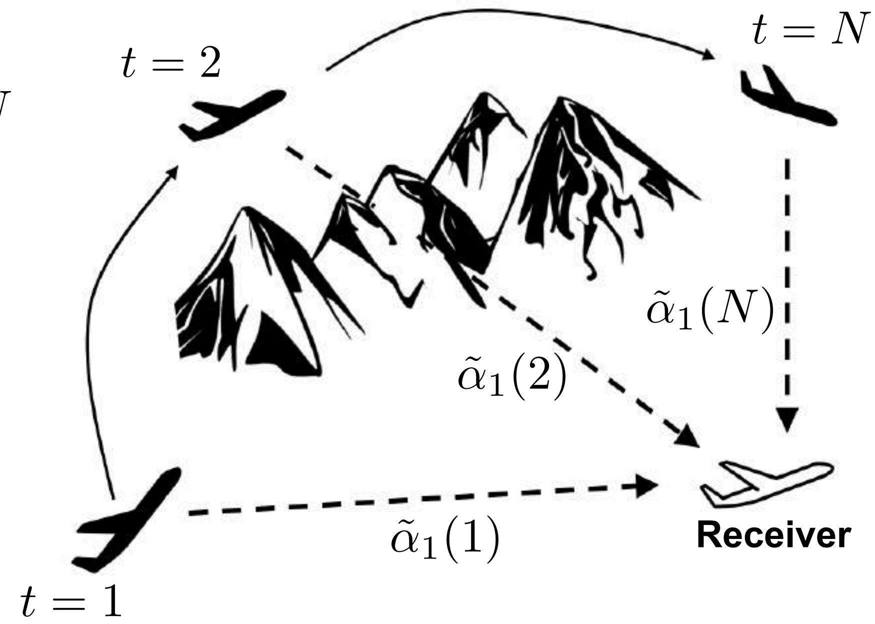


Model Misspecification – Dynamic Channel

Assumed Model: Constant channel gains



True Model: Time-varying channel gains



Model Misspecification – Dynamic Channel

- **Assumed model** $\Longrightarrow$ $\alpha_1(t) = \alpha_1$ **(Constant Channel Gain)**

$$\tilde{\mathbf{x}} \sim \mathcal{CN}(\boldsymbol{\mu}(\boldsymbol{\theta}), \tilde{\mathbf{R}}) = p_{\tilde{\mathbf{x}}|\boldsymbol{\theta}} \Longrightarrow \boldsymbol{\mu}(\boldsymbol{\theta}) = \begin{bmatrix} \alpha_1 s_1(1)\mathbf{a}(\phi_1) \\ \vdots \\ \alpha_1 s_1(N)\mathbf{a}(\phi_1) \end{bmatrix}$$

- **True Model** $\Longrightarrow$ **Channel Gains Vary with time**

$$\tilde{\mathbf{x}} \sim \mathcal{CN}(\mathbf{d}, \tilde{\mathbf{R}}) = q_{\tilde{\mathbf{x}}} \Longrightarrow \mathbf{d} = \begin{bmatrix} \tilde{\alpha}_1(1) s_1(1)\mathbf{a}(\phi_1) \\ \vdots \\ \tilde{\alpha}_1(N) s_1(N)\mathbf{a}(\phi_1) \end{bmatrix}$$

$$\boxed{\boldsymbol{\theta}_A = \arg\min_{\boldsymbol{\theta}} D(q_{\tilde{\mathbf{x}}} \,||\, p_{\tilde{\mathbf{x}}|\boldsymbol{\theta}})}$$

Ira A. Fulton Schools of Engineering
Arizona State University

Model Misspecification – Dynamic Channel

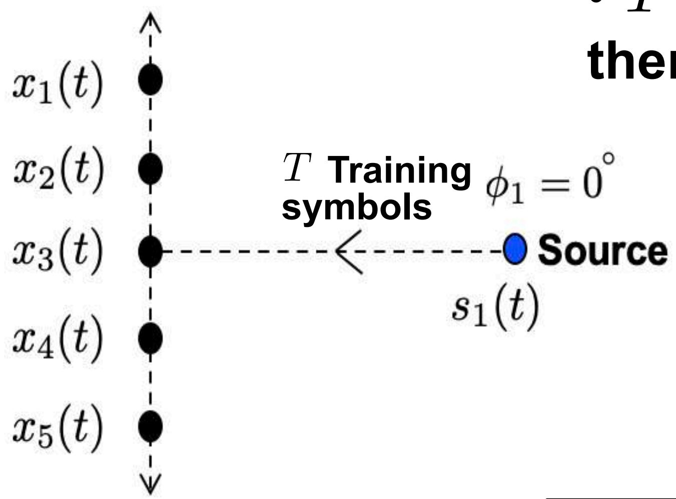
- **It can be shown that**

$$\boldsymbol{\theta}_A = \left[s_1(1), \ldots, s_1(N), s_1^*(1), \ldots, s_1^*(N), \alpha_{1,A}, \alpha_{1,A}^*, \phi_{1,A}\right]^T$$

- $$\phi_{1,A} = \arg\max_{\phi} \frac{|\mathbf{a}^H(\phi)\tilde{\mathbf{R}}^{-1}\mathbf{a}^H(\phi_1)|^2}{\mathbf{a}^H(\phi)\tilde{\mathbf{R}}^{-1}\mathbf{a}^H(\phi)} \cdot \frac{\left|\sum_{i=1}^{T} |s_1(i)|^2 \tilde{\alpha}_1(i)\right|^2}{\sum_{i=1}^{T} |s_1(i)|^2}$$

- $$\alpha_{1,A} = \frac{\mathbf{a}^H(\phi_{1,A})\tilde{\mathbf{R}}^{-1}\mathbf{a}^H(\phi_1)}{\mathbf{a}^H(\phi_{1,A})\tilde{\mathbf{R}}^{-1}\mathbf{a}^H(\phi_{1,A})} \cdot \frac{\left|\sum_{i=1}^{T} |s_1(i)|^2 \tilde{\alpha}_1(i)\right|^2}{\sum_{i=1}^{T} |s_1(i)|^2}$$

Semi-Blind Constraint for Communications Model



- T **training symbols available for source** j**,** **then the** $m = 2T$ **complex constraints on** θ **are:**

$$\left.\begin{aligned} f_{2i-1}(\boldsymbol{\theta}) &= s_j(i) - s_{ji} = 0 \\ f_{2i}(\boldsymbol{\theta}) &= s_j^*(i) - s_{ji}^* = 0 \end{aligned}\right\} i = 1, \ldots, T$$

- **Gradient Matrix**

$$\mathbf{F}_j(\boldsymbol{\theta}) = \left[\mathbf{I}_T \otimes \mathbf{E}_j, \quad \mathbf{0}_{2T\times(N-T)k}, \quad \mathbf{I}_T \otimes \mathbf{G}_j\right]$$

where $\mathbf{E}_j = \begin{bmatrix} \mathbf{0}_{1\times k} \\ \mathbf{e}_j^T \end{bmatrix}$**,** $\mathbf{G}_j = \begin{bmatrix} \mathbf{e}_j^T \\ \mathbf{0}_{1\times k} \end{bmatrix}$ **and** $\mathbf{e}_j^T = [0, \ldots, 0, \underset{\uparrow\, j^{th} \text{ position}}{1}, 0, \ldots 0]_{1\times k}$ **.**

- **Null matrix** $\mathbf{U}$ **can be computed such that** $\mathbf{F}(\boldsymbol{\theta})\mathbf{U} = \mathbf{0}_{m\times(2Nk+3k-m)}$**.**

Ira A. Fulton Schools of Engineering
Arizona State University

Constrained Misspecified Cramer-Rao Bound

- **Thus we have found θ_A and U.**
- **$\mathrm{J}(\boldsymbol{\theta})$ and average Hessian $\mathrm{C}(\theta)$ derived in the paper.**
- **Constrained Misspecified CRB:**

$$\mathbb{E}\left[\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_A\right)\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_A\right)^H\right] \geq \mathbf{U}(\mathbf{U}^H\mathbf{C}(\boldsymbol{\theta}_A)\mathbf{U})^{-1}\mathbf{U}^H \times \mathbf{J}(\boldsymbol{\theta}_A)\mathbf{U}(\mathbf{U}^H\mathbf{C}(\boldsymbol{\theta}_A)\mathbf{U})^{-1}\mathbf{U}^H$$

Numerical Results

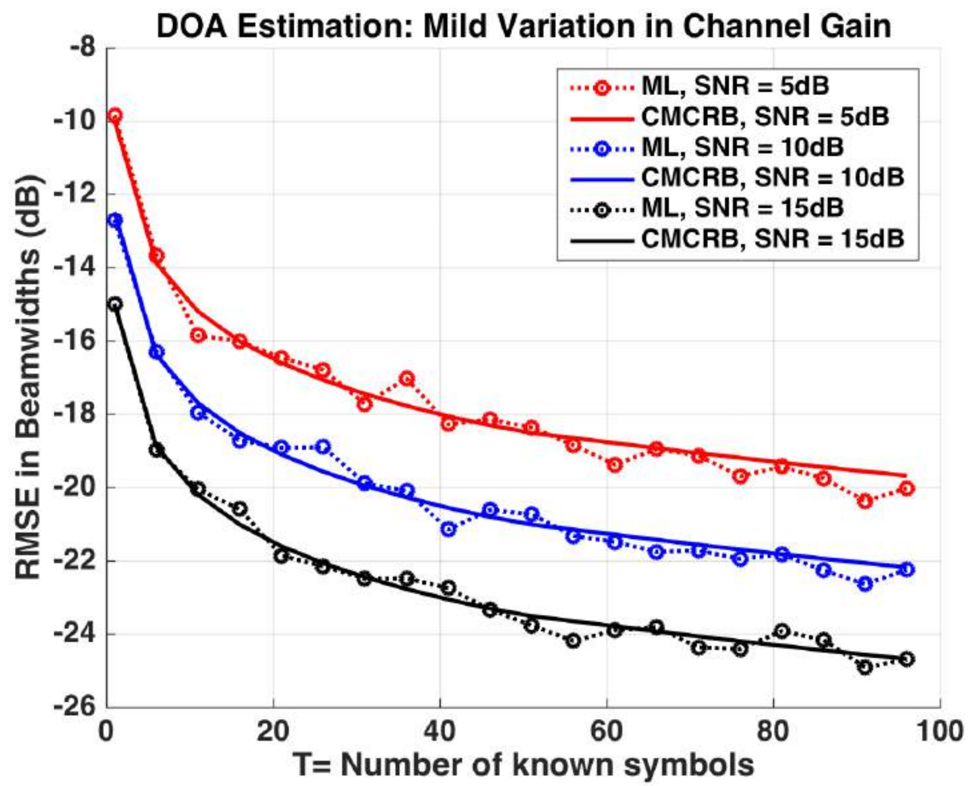
- **Assumed model: α_1 constant.**
- **True Model:**

$$\tilde{\alpha}_1(t) = \begin{cases} e^{j100^\circ}, & t = 1, 2, \ldots, 50 \\ e^{j113^\circ}, & t = 51, 52, \ldots, 100. \end{cases}$$

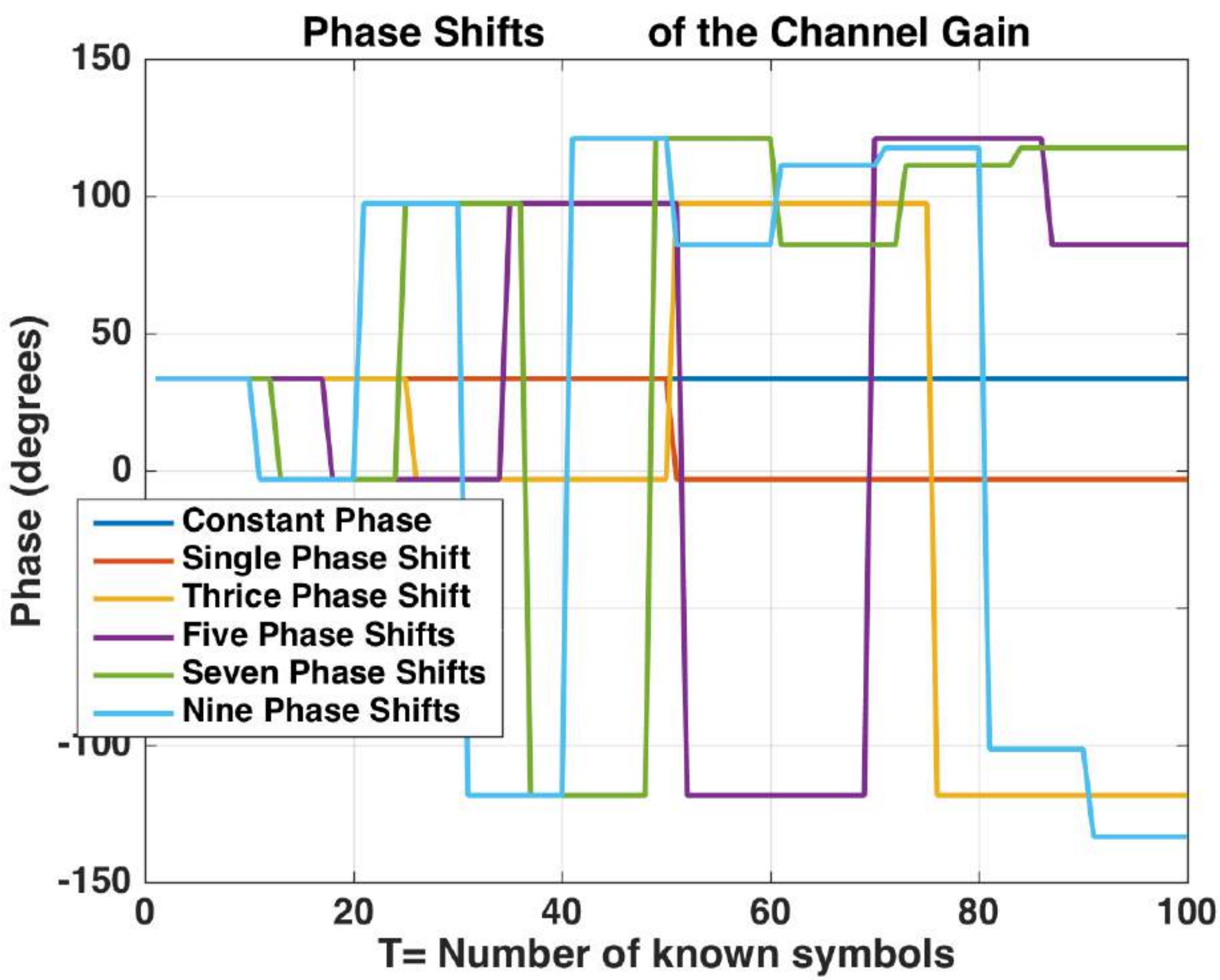
- **Maximum Likelihood Estimate:**

$$\hat{\phi}_{1,ML} = \arg\max_{\phi} \frac{\left|\mathbf{v}^H(\phi)\tilde{\mathbf{R}}^{-1}\tilde{\mathbf{x}}\right|^2}{\mathbf{v}^H(\phi)\tilde{\mathbf{R}}^{-1}\mathbf{v}(\phi)}$$

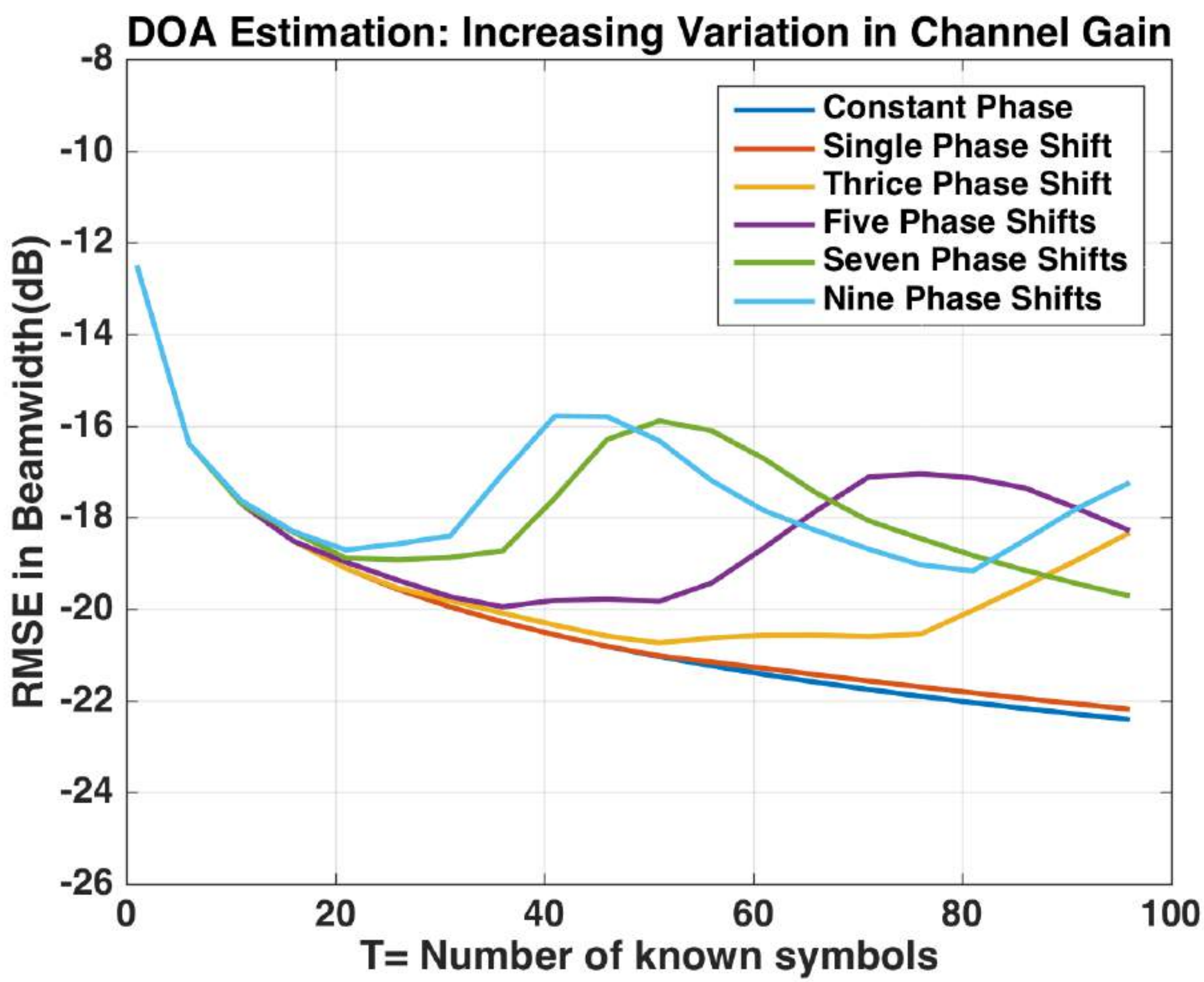
$$\mathbf{v}(\phi) = \begin{bmatrix} s_1(1)\mathbf{a}(\phi) \\ \vdots \\ s_1(N)\mathbf{a}(\phi) \end{bmatrix}$$



Numerical Results



Numerical Results



Ira A. Fulton Schools of Engineering
Arizona State University

Summary

- **Parameter estimation bounds for a communication link under model misspecification.**

- **Constrained Misspecified Cramer Rao Bound (CMCRB).**

- **Dynamic Channel – Bounds on angle of arrival when channel gains vary with time but are assumed to be constant.**

- **Effect on bounds when dynamic nature of channel is varied.**

- **Different forms of dynamic channels can be considered.**

Outline

- Introduction
- Akshay's discussion
- **Closing remarks**

Closing Remarks

- Initial results show encouraging agreement, and more complex channels can be considered
  - We welcome input from others here…
    - Perhaps inform of important use cases to focus on
    - Perhaps inform of models that have been successfully employed, etc.
  - Perhaps inform what can be assumed about TX/RX hardware?
    - e.g. # antennas, bandwidth, modulation schemes, etc.

- MCRB can help identify most essential aspects of modeling, e.g.
  - Determine which dynamics must be properly captured
  - versus which dynamics are less important

- CRB / MCRB can provide helpful reference points for performance comparisons

Backups

Regularity Conditions: Allow Reverse Order of Integration and Differentiation

- **By linearity of integration we have**

$$E\left\{(\widehat{\theta}-\theta)\frac{\partial \ln p}{\partial \theta}\right\} = \int (\widehat{\theta}-\theta)\frac{\partial \ln p}{\partial \theta}\, p d\mathbf{x}$$

$$= \int \widehat{\theta}\frac{\partial \ln p}{\partial \theta}\, p d\mathbf{x} - \theta \int \frac{\partial \ln p}{\partial \theta}\, p d\mathbf{x}$$

- **Order reversal of integration / differentiation simplifies first term:**

$$\int \widehat{\theta}\frac{\partial \ln p}{\partial \theta}\, p d\mathbf{x} = \int \widehat{\theta} \cdot \frac{1}{p}\frac{\partial p}{\partial \theta} \cdot p d\mathbf{x}$$

$$= \int \widehat{\theta} \cdot \frac{\partial p}{\partial \theta} d\mathbf{x} = \frac{\partial}{\partial \theta}\int \widehat{\theta} \cdot p d\mathbf{x} = \frac{\partial \theta}{\partial \theta} = 1$$

- **Lastly, order reversal shows that second term vanishes:**

$$\int \frac{\partial \ln p}{\partial \theta} \cdot p d\mathbf{x} = \int \frac{1}{p}\frac{\partial p}{\partial \theta} \cdot p d\mathbf{x}$$

$$= \int \frac{\partial p}{\partial \theta} d\mathbf{x} = \frac{\partial}{\partial \theta}\int p d\mathbf{x} = \frac{\partial}{\partial \theta}(1) = 0$$

Return