

Uncertainty quantification in the time domain pipeline leak detection

Alireza Keramat

June 2017

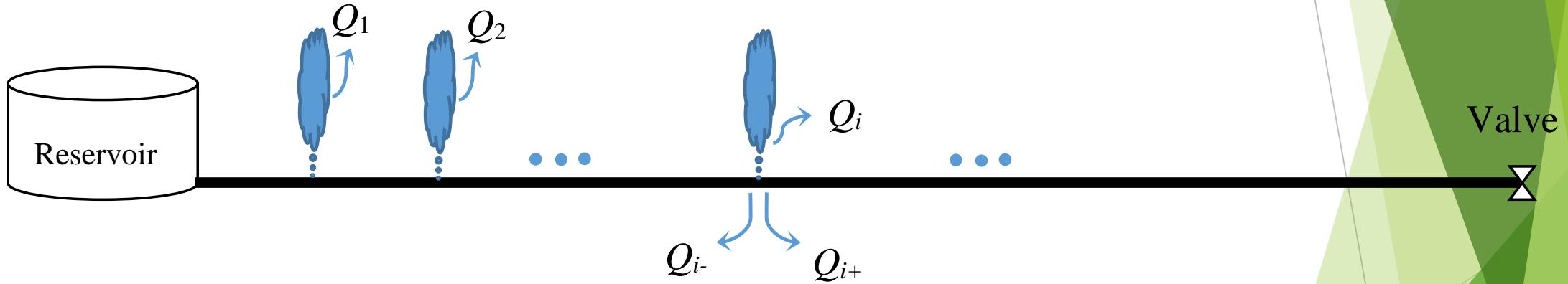
Introduction

- ▶ Time domain model for leak detection in a noisy environment
- ▶ Cramer-Rao lower bound (CRLB)
- ▶ Direct differentiation method (DDM) for numerical computation of CRLB
- ▶ Case study, results and discussion
 - ✓ **Time closure**
 - ✓ **Sample size**
 - ✓ **Noise level**
 - ✓ **Correlation between adjacent leaks, resolution in localization**

Leak detection in the time domain

$$\mathbf{h}_m = \mathbf{h}(x_{Li}, A_{ei}) + \mathbf{n}, \quad \mathbf{h}_m = (h_{m,1}, \dots, h_{m,M})^T, \quad \mathbf{h} = (h_1, \dots, h_M)^T, \quad \mathbf{n} = (n_1, \dots, n_M)^T$$

- Water hammer model (continuity and momentum equations)
- Method of characteristics



$$\mathbf{h}(x_{Li}, A_{ei}) \approx \mathbf{h}(x_j, A_{ej}), \quad j = 1, \dots, N \quad A_{ej} = A_{ej}(x_j)$$

$$\mathbf{h}_m = \mathbf{h}(A_{ej}) + \mathbf{n}, \quad j = 1, \dots, N$$

Cramer Rao Lower Bound

Provides a lower bound for the variance of the fitted parameters

- Unbiased estimator
- The model is complete and correct

A measure for maximum precision

$$\text{var}(\hat{\theta}) \geq \frac{1}{I(\theta)}$$

where the Fisher information $I(\theta)$ is defined by

$$I(\theta) = E \left[\left(\frac{\partial \ell(x; \theta)}{\partial \theta} \right)^2 \right] = -E \left[\frac{\partial^2 \ell(x; \theta)}{\partial \theta^2} \right]$$

Assumptions:
wave speed,
friction factor,
and closure
time and
pattern $\tau(t)$ are
deterministic,
measurements
are random

Cramer Rao Lower Bound in the Inverse Transient Analysis (ITA)

$$p(\mathbf{h}_m, \mathbf{A}_e) = \frac{1}{\sqrt{(2\pi)^M \det(\Sigma)}} \exp\left(-\frac{1}{2} (\mathbf{h}_m - \mathbf{h})^T \Sigma^{-1} (\mathbf{h}_m - \mathbf{h})\right).$$

$$\ln L(\mathbf{A}_e; \mathbf{h}_m) = -\frac{M}{2} \ln(2\pi) - M \ln \sigma - \frac{1}{2\sigma^2} \|(\mathbf{h}_m - \mathbf{h})\|^2$$

CRLB: $\text{cov}(\hat{\mathbf{A}}_e) \geq \mathbf{I}(\mathbf{A}_e)^{-1}$,

$$I_{i,j} = -\mathbf{E}\left[\frac{\partial^2}{\partial A_{ei} \partial A_{ej}} \ln L(\mathbf{A}_e; \mathbf{h}_m)\right]$$

Cramer Rao Lower Bound in the Inverse Transient Analysis (ITA)

Each element of the Fisher Information Matrix (FIM):

$$I_{i,j} = -\mathbf{E} \left[\frac{\partial^2}{\partial A_{ei} \partial A_{ej}} \ln L(\mathbf{A}_e; \mathbf{h}_m) \right] = \frac{1}{\sigma^2} \frac{\partial \mathbf{h}^T}{\partial A_{ei}} \frac{\partial \mathbf{h}}{\partial A_{ej}}.$$

\mathbf{h} contains a, f , measurement location, time and pattern of valve closure, etc.

$$\text{cov}(\hat{\mathbf{A}}_e) \geq \mathbf{I}(\mathbf{A}_e)^{-1}, \quad \text{cov}(\hat{\mathbf{A}}_e) = \mathbf{E} \left[(\mathbf{A}_e - \hat{\mathbf{A}}_e)(\mathbf{A}_e - \hat{\mathbf{A}}_e)^T \right] \geq \mathbf{I}(\mathbf{A}_e)^{-1}$$

Partial derivatives are calculated using Characteristic equations.

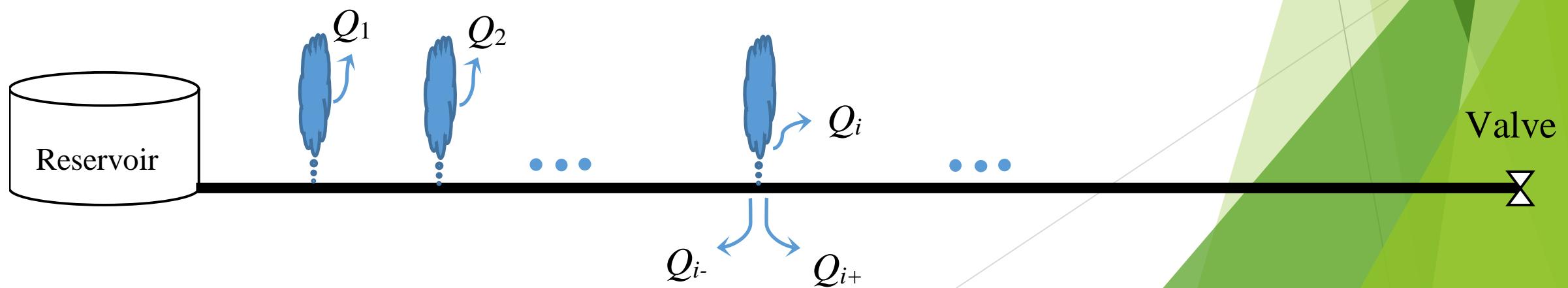
Direct Differentiation Method (DDM) to compute sensitivities

Steady state

$$\frac{\partial h_1}{\partial A_{ei}} = 0, \quad \frac{\partial Q_{1+}}{\partial A_{ei}} = 0,$$

$$\frac{\partial Q_{k,+}}{\partial A_{ei}} = \frac{\partial Q_{k,-}}{\partial A_{ei}} - \sqrt{2g(h_k - h_o)} \left(\frac{\partial A_{ek}}{\partial A_{ei}} + \frac{A_{ek}}{2h_k} \frac{\partial h_k}{\partial A_{ei}} \right)$$

$$\frac{\partial Q_{k,-}}{\partial A_{ei}} = \frac{\partial Q_{k-1,+}}{\partial A_{ei}}, \quad \frac{\partial h_k}{\partial A_{ei}} = \frac{\partial h_{k-1}}{\partial A_{ei}} - \frac{f \Delta x Q_{k-1,+}}{g D A^2} \frac{\partial Q_{k-1,+}}{\partial A_{ei}}, \quad k = 2, 3, \dots, N$$



Direct Differentiation Method (DDM) to compute sensitivities

Transient state

$$C^+ : \quad Q_{k-} = -C_{ap} h_k + C_p, \quad C^- : \quad Q_{k+} = C_{an} h_k + C_n, \quad Q_{1k} - Q_{2k} = A_{ek} \sqrt{2g(h_k - h_o)},$$

$$h_k = -\frac{A_{ek}\alpha^{0.5} - gA_{ek}^2 + C_n C_{an} + C_n C_{ap} - C_p C_{an} - C_p C_{ap}}{(C_{an} + C_{ap})^2}, \text{ with}$$

$$\alpha = -g(2C_{an}^2 h_o - gA_{ek}^2 + 2C_{ap}^2 h_o + 2C_n C_{an} + 2C_n C_{ap} - 2C_p C_{an} - C_p C_{ap} + 4C_{an} C_{ap} h_o).$$

$$\frac{\partial h_k}{\partial A_{ei}} = \frac{\partial h_k}{\partial C_n} \frac{\partial C_n}{\partial A_{ei}} + \frac{\partial h_k}{\partial C_p} \frac{\partial C_p}{\partial A_{ei}} + \frac{\partial h_k}{\partial C_{ap}} \frac{\partial C_{ap}}{\partial A_{ei}} + \frac{\partial h_k}{\partial C_{an}} \frac{\partial C_{an}}{\partial A_{ei}} + \frac{\partial h_k}{\partial A_{ei}}$$

Numerical case study



| | | | |
|-------------------------------|----------|------------------------------------|--------------------|
| Pipe length | 1000 m | Number of MOC points | 20 |
| Inner diameter | 0.2 m | Mesh size | 50 m |
| Steady velocity (before leak) | 0.2 m/s | Leak distance from reservoir | 500 |
| Wave speed | 1310 m/s | Effective leak sizes ($C_d A_L$) | 20 mm ² |

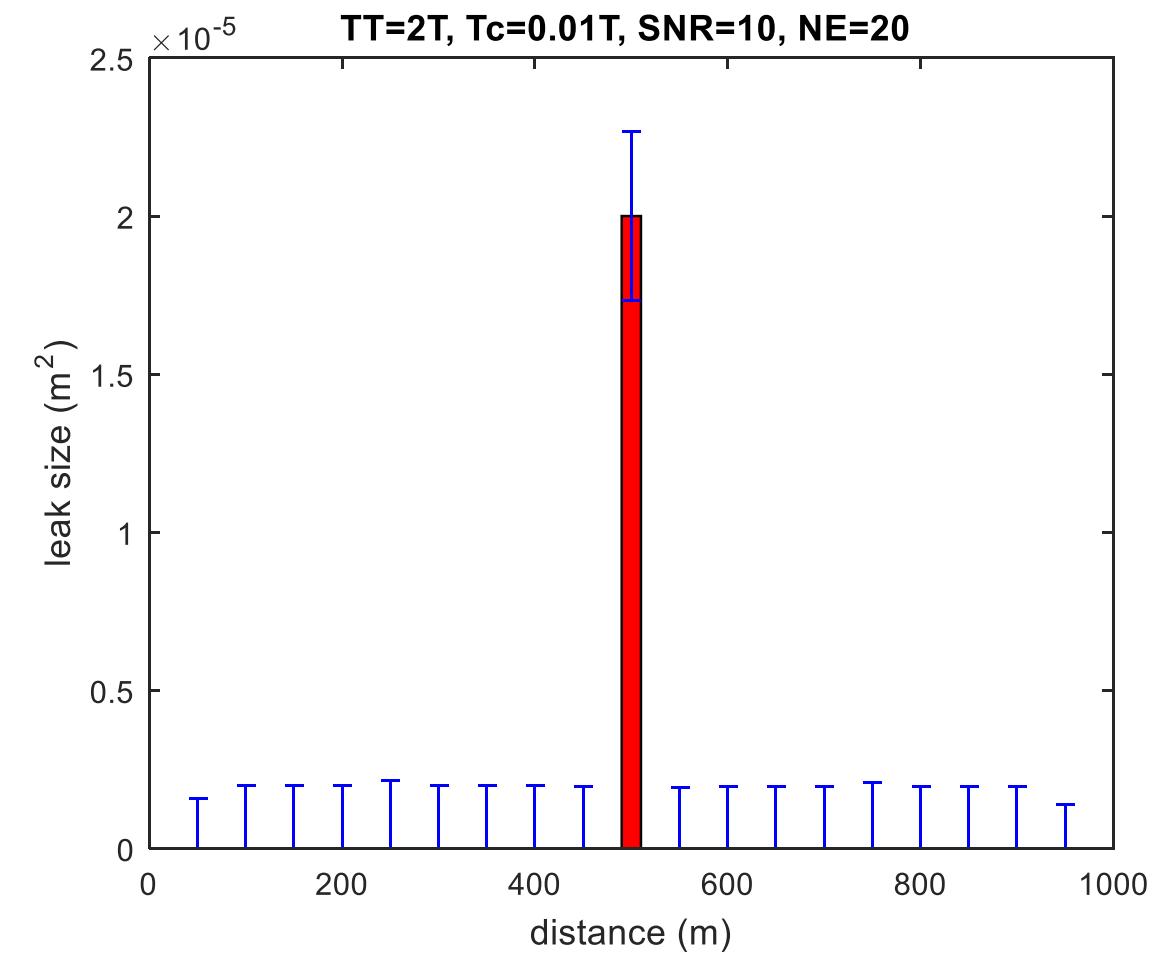
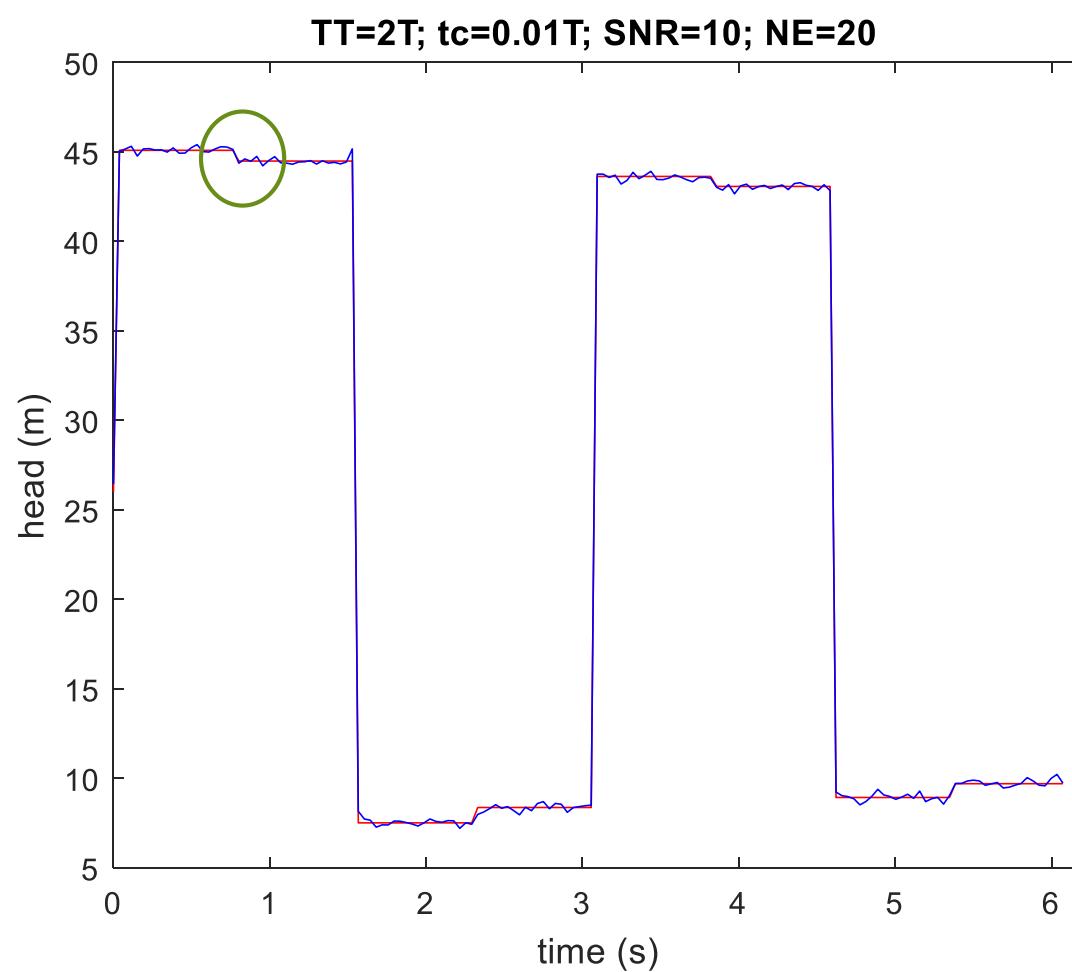
Gaussian white noise

$$\mathbf{n} = \sigma \boldsymbol{\xi}, \quad \sigma = 10^{-\frac{\text{SNR}}{20}} \Delta h_L \quad \text{or} \quad \text{SNR} = 10 \log_{10} \left(\frac{|\Delta h_L|}{\sigma} \right)^2$$

Total measurement = $2T$

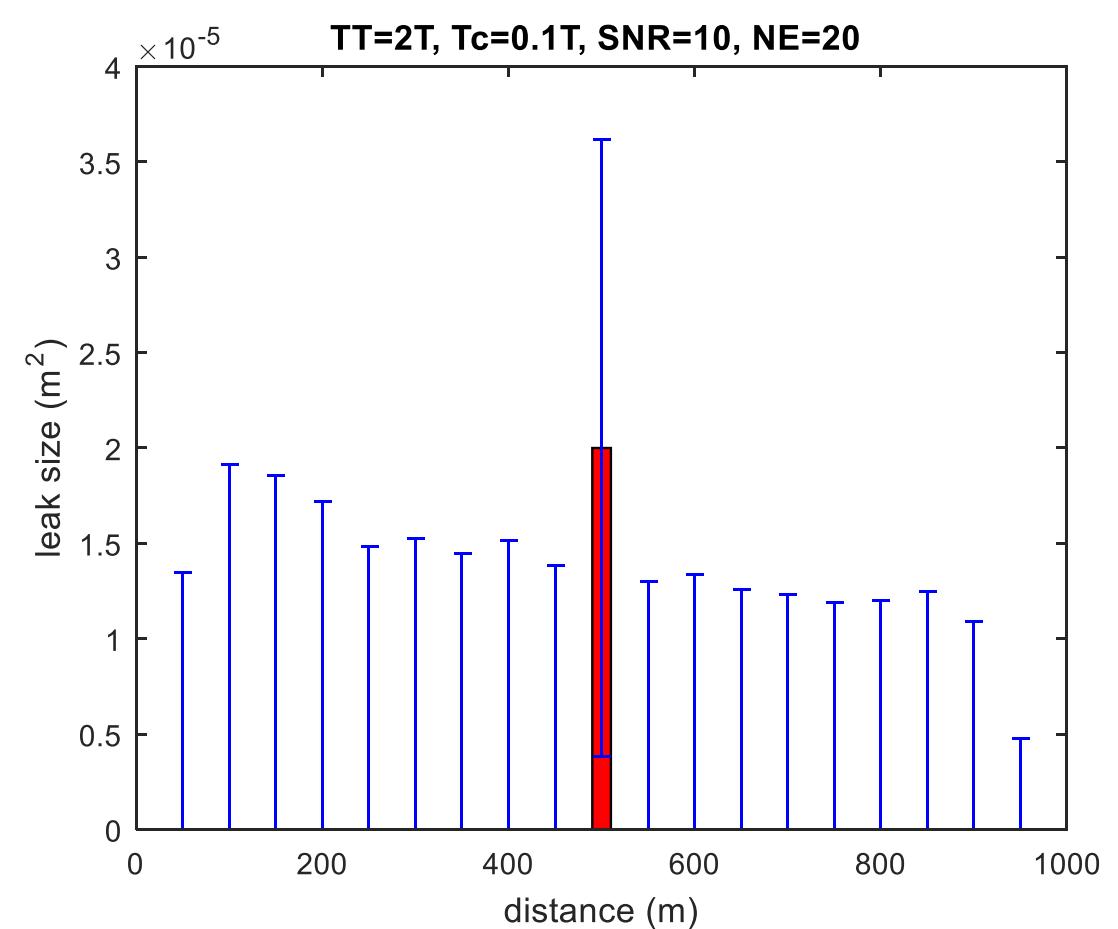
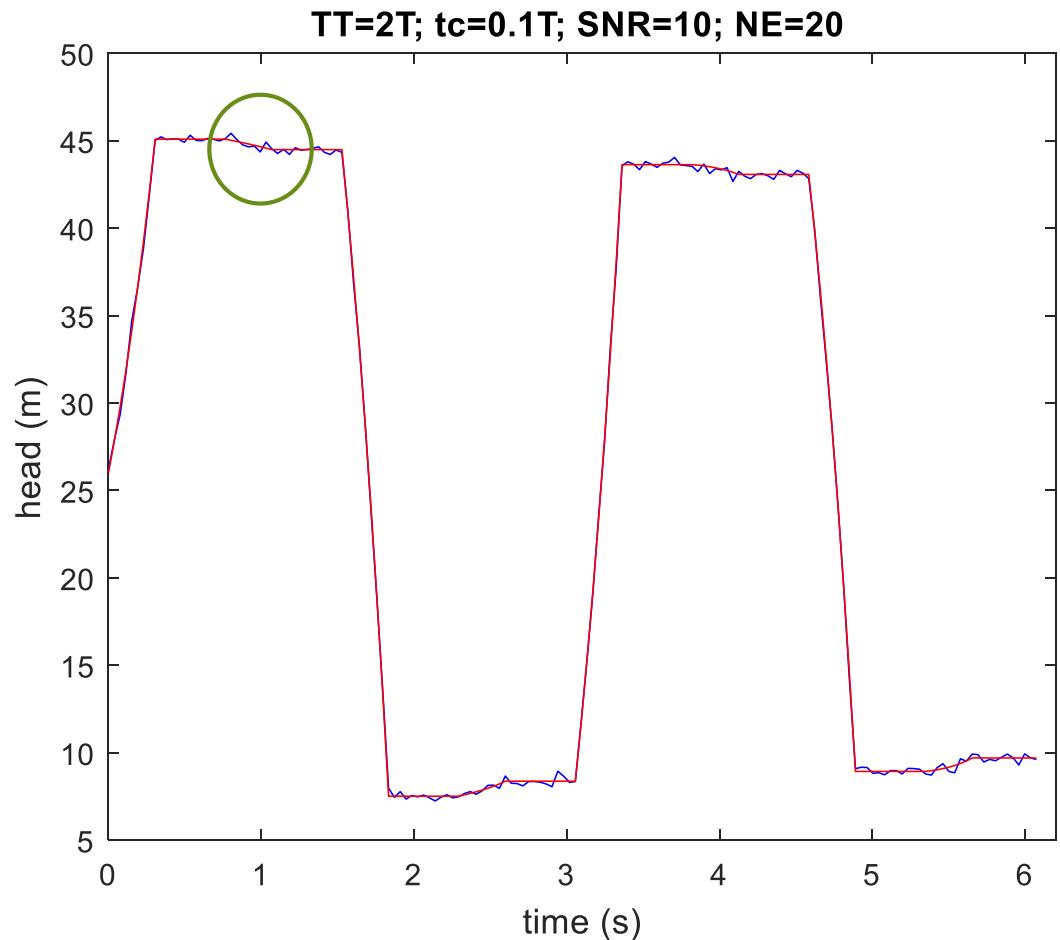
T closure= $0.01T$

SNR=10 dx=50m



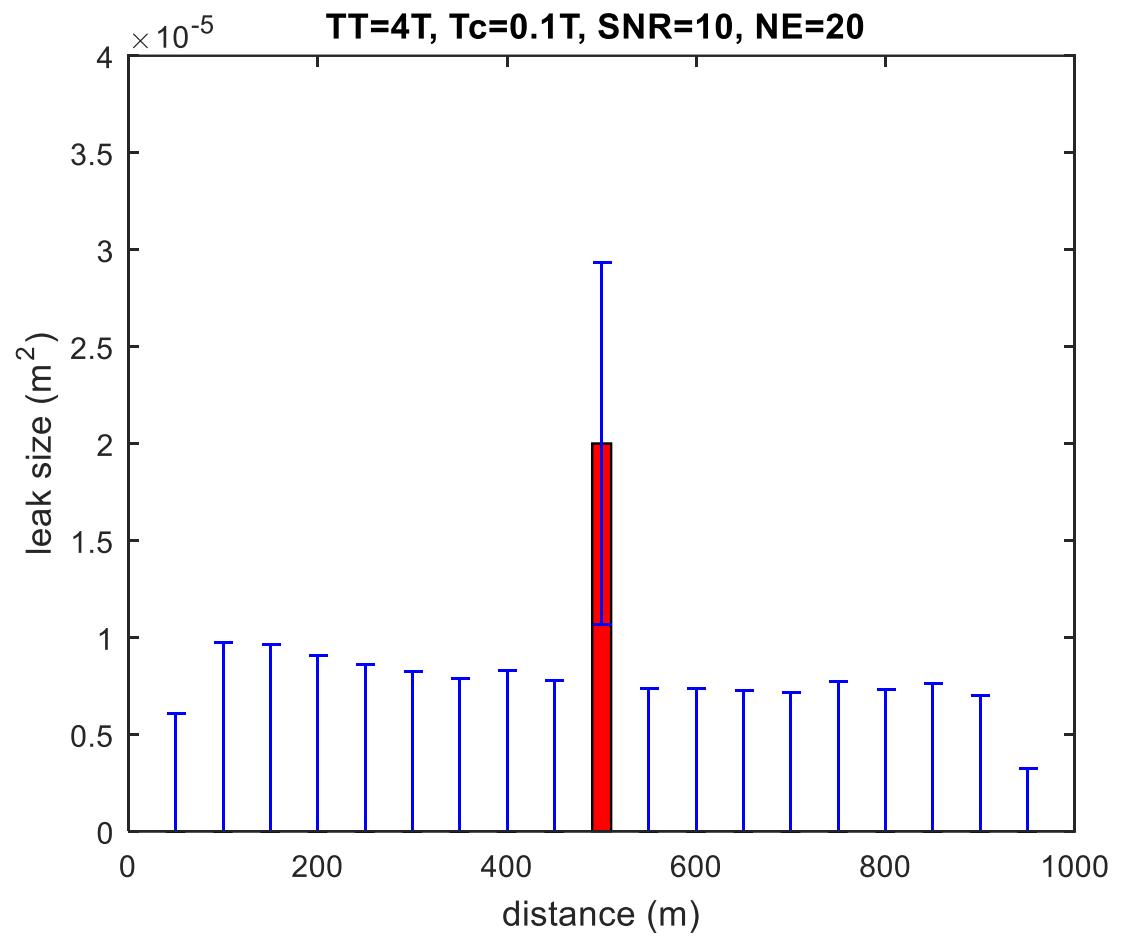
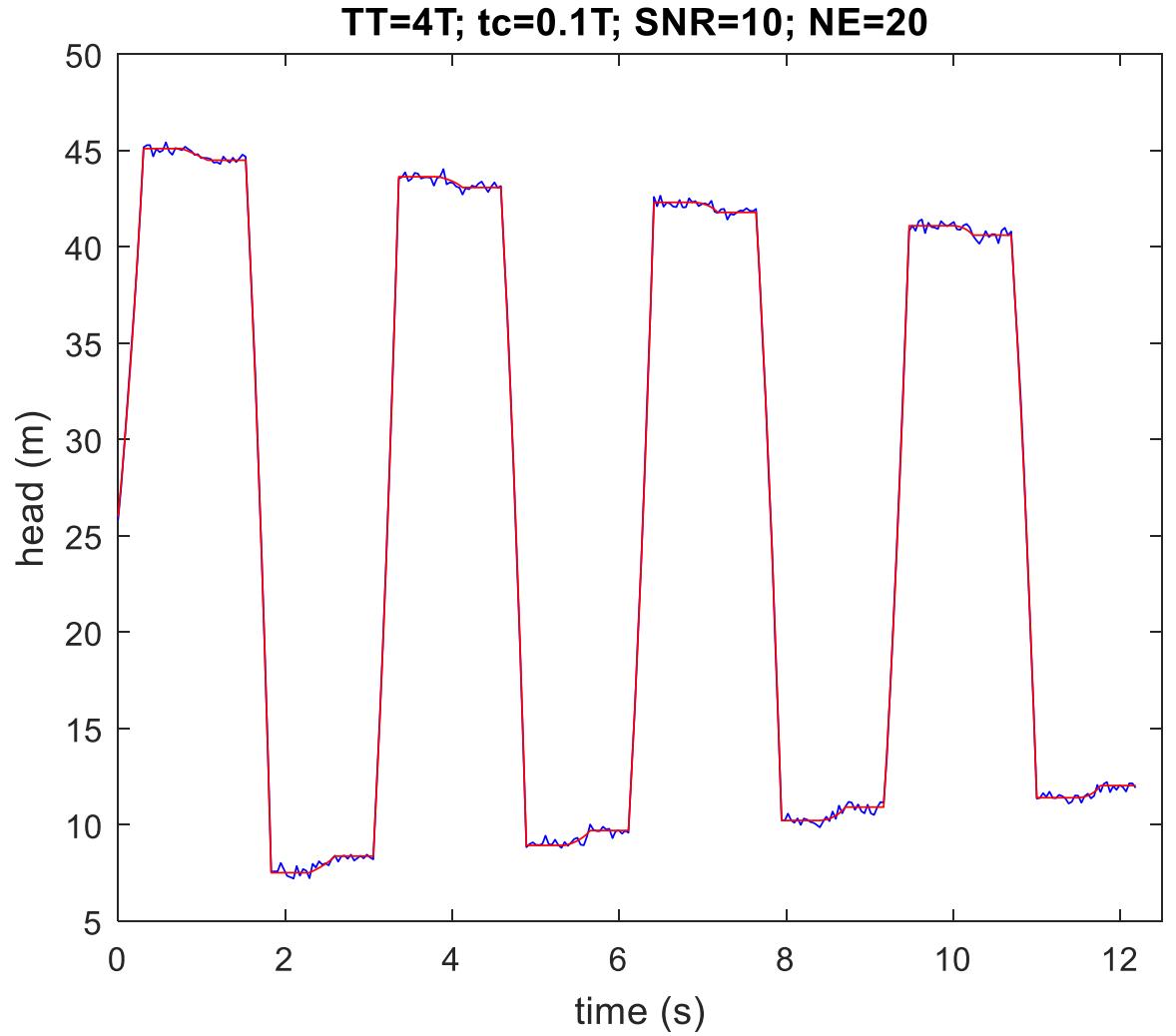
Increase time closure to 0.1T wave front = $a \cdot T_c = 400$ m

Inaccurate size detection and localization

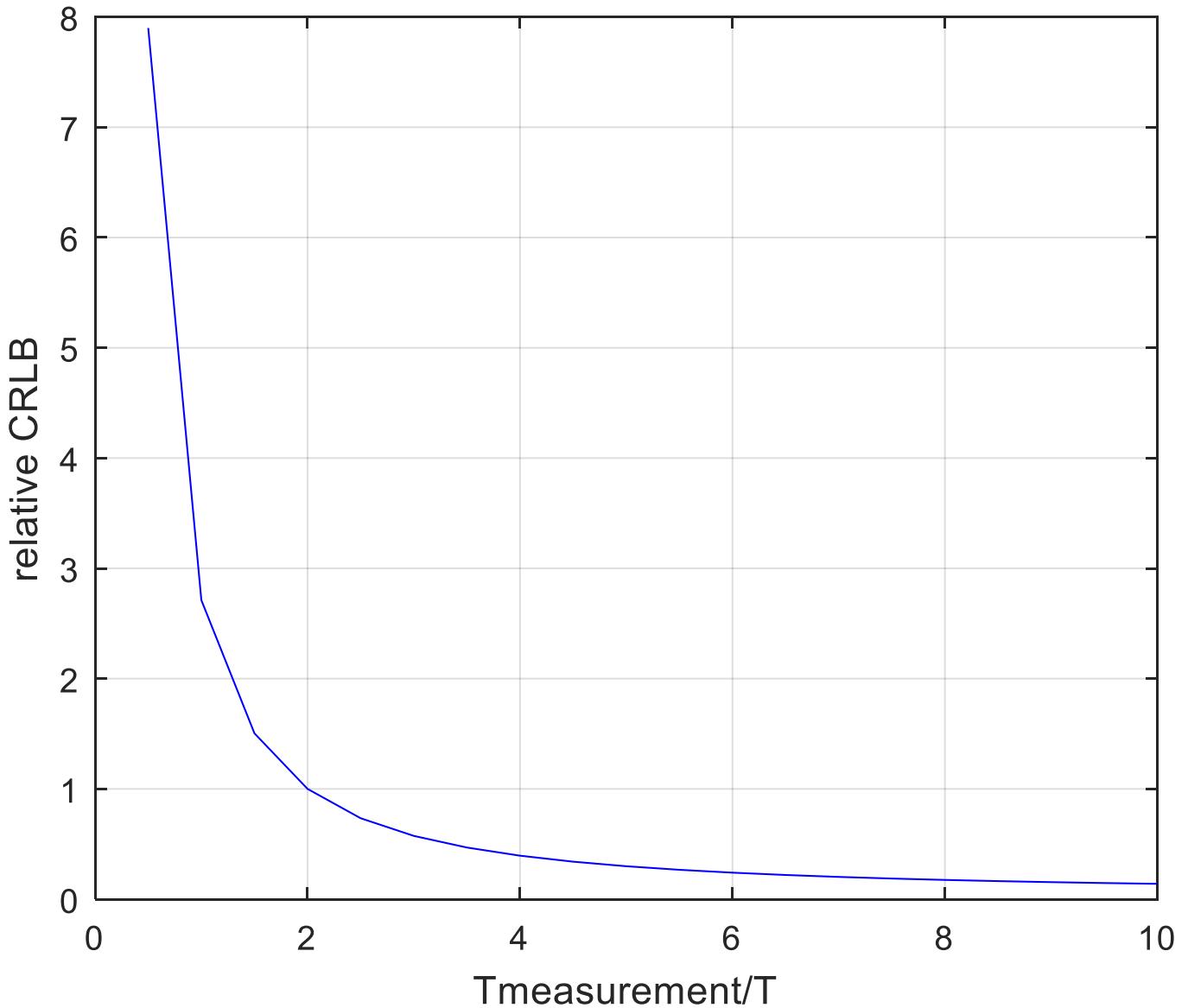
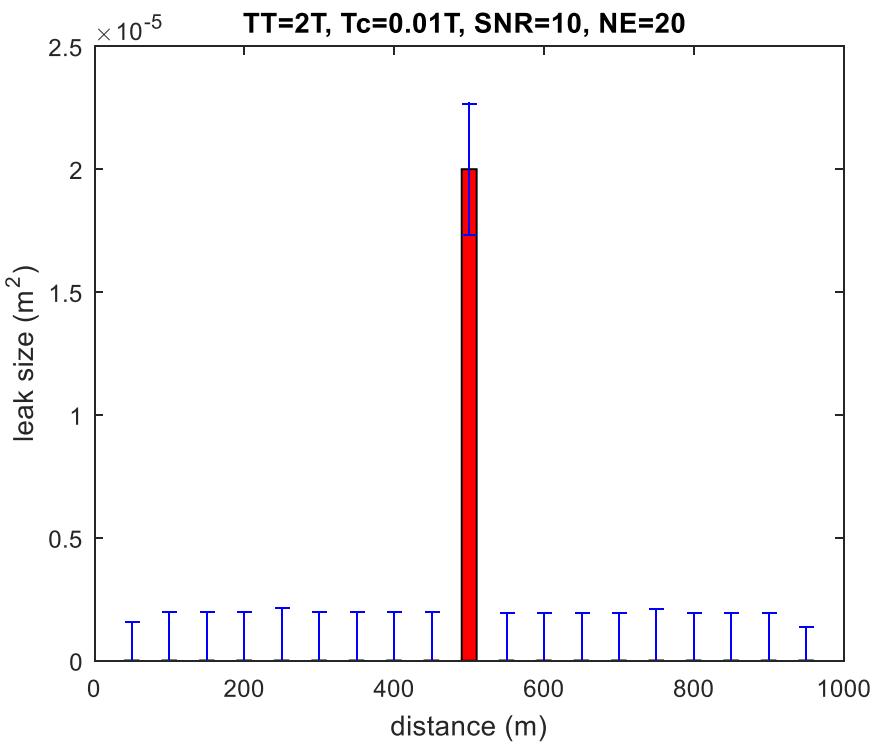


Measurements up to 4 periods

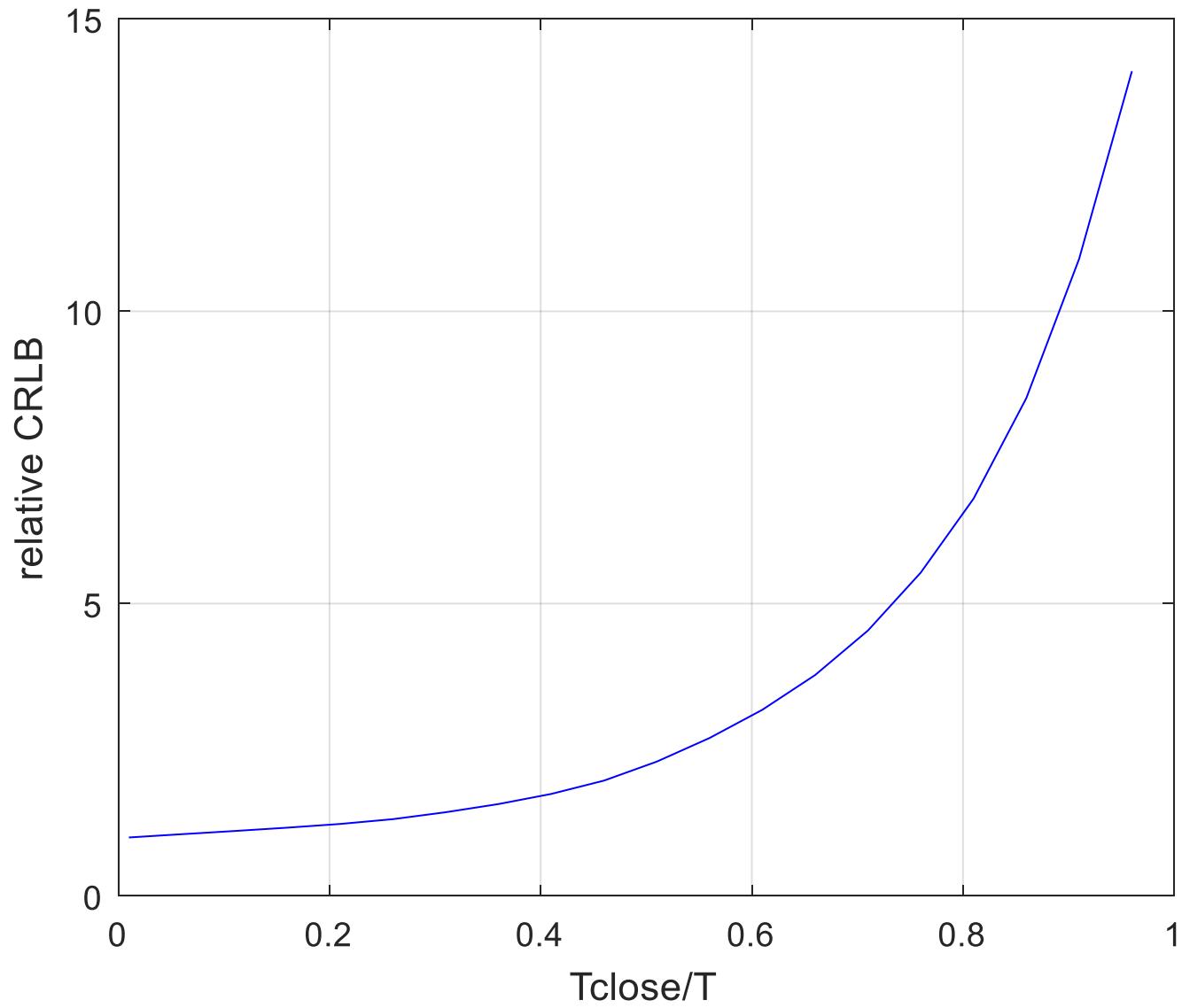
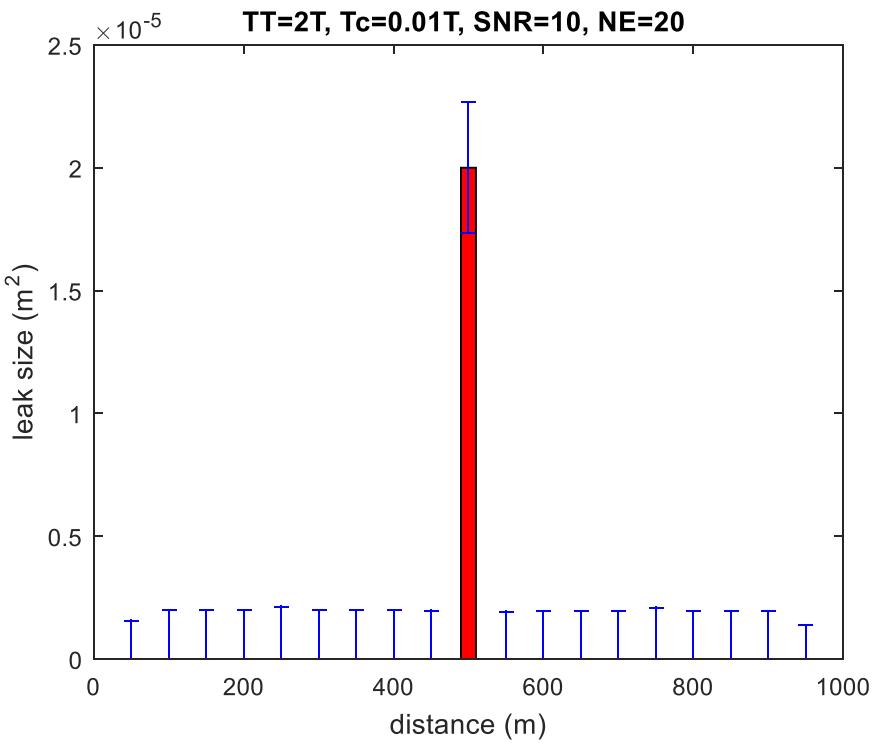
Inaccurate size detection and but good localization in 50 m



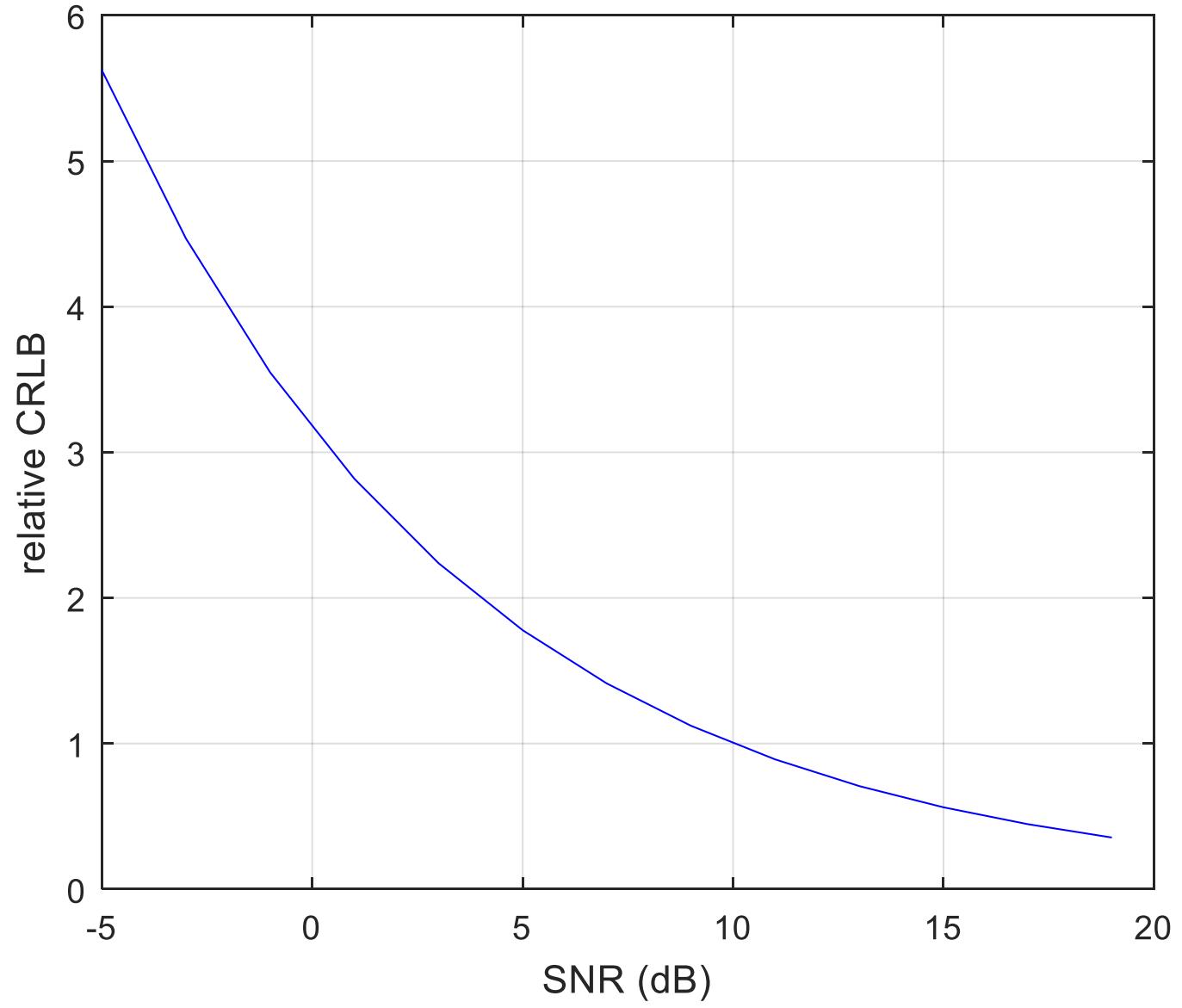
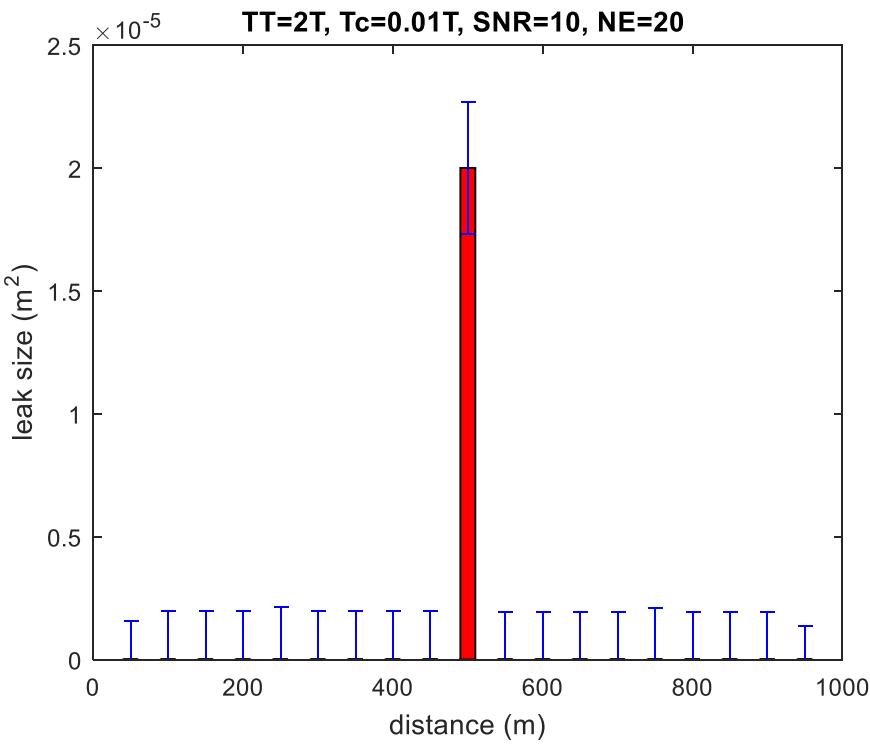
Increasing the number of measurements



Increasing the closure time



Increasing SNR

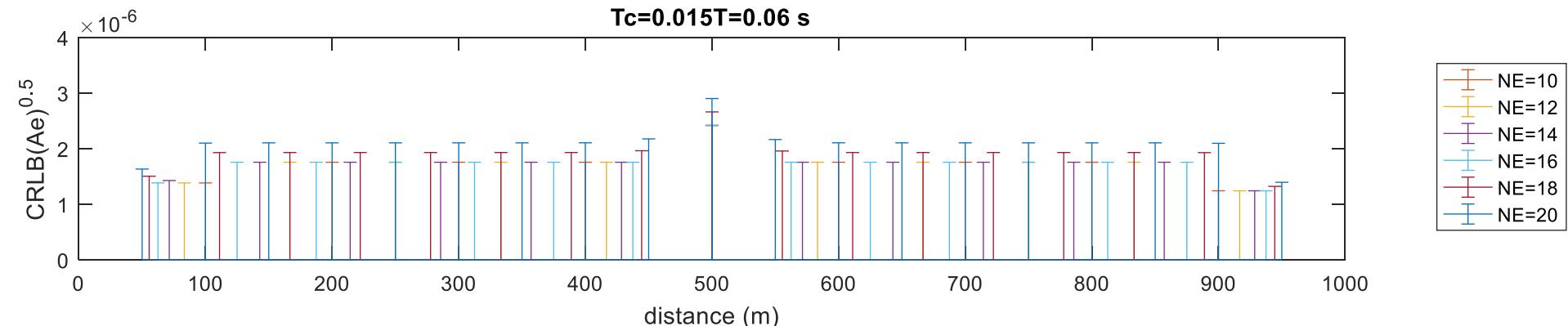
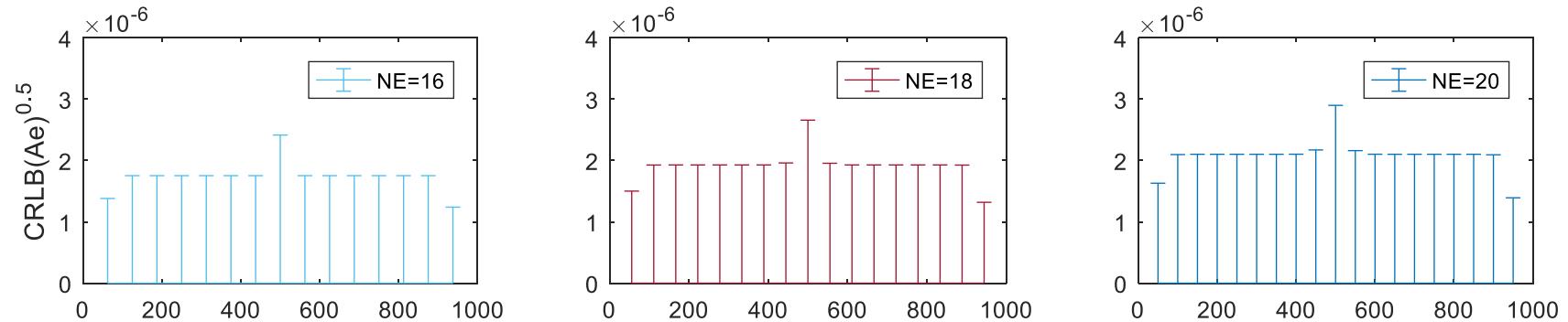
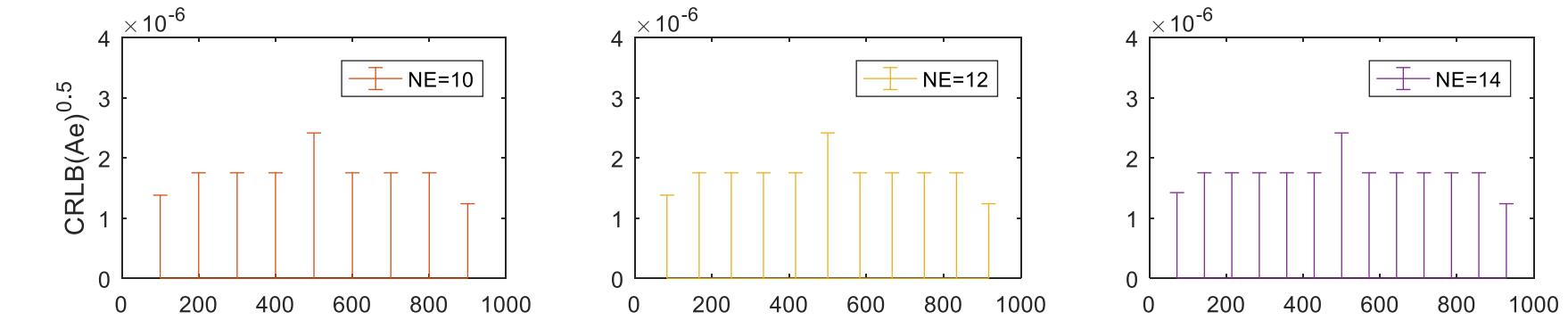


Study the number of parameters (distance between leaks)

CRLB does
not change if
 $\Delta x_{MOC} > a \cdot T_c$

$T_c = 0.06\text{s}$
 $\lambda = 60\text{m}$

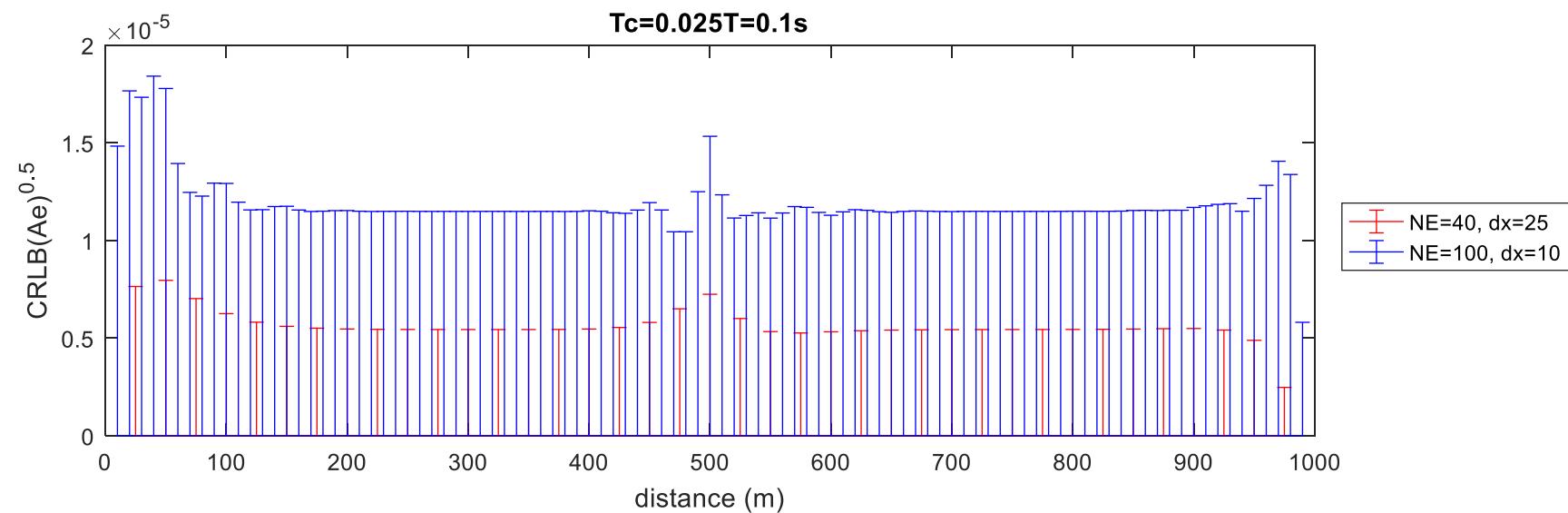
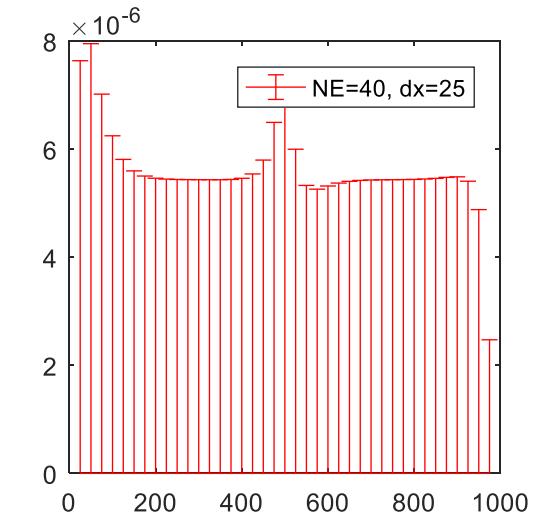
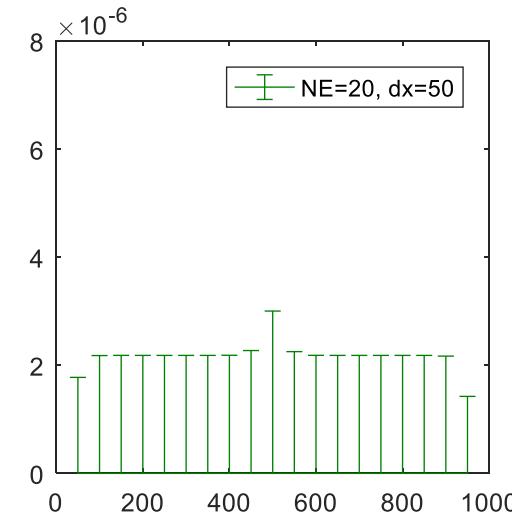
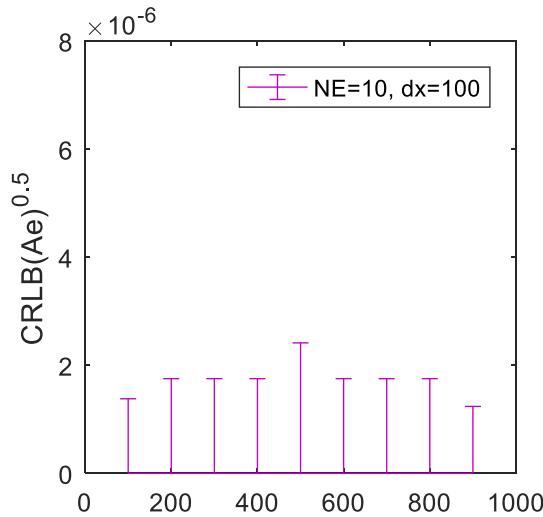
$\Delta x_{MOC} > a \cdot T_c$
 $NE < 1000/60 = 16$



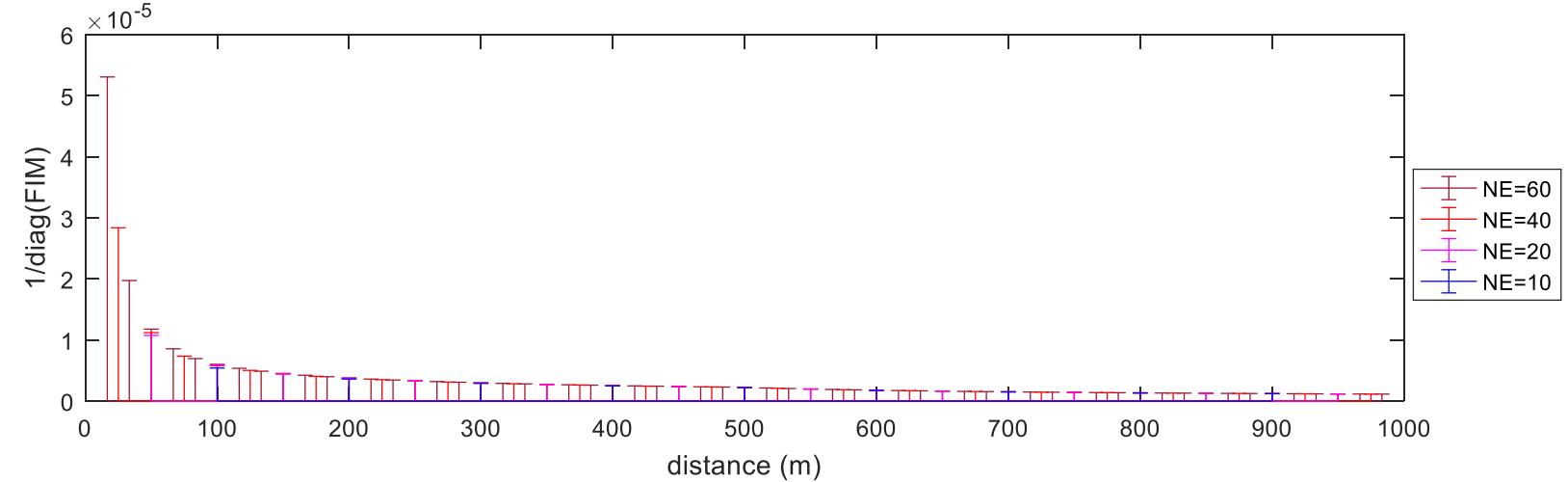
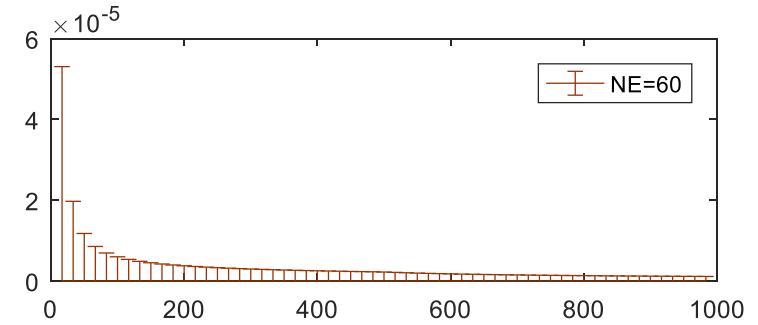
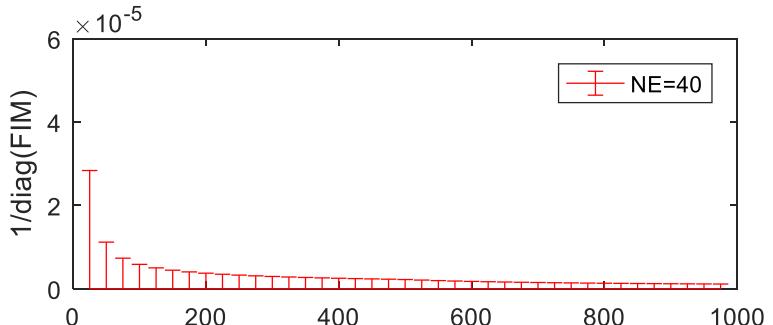
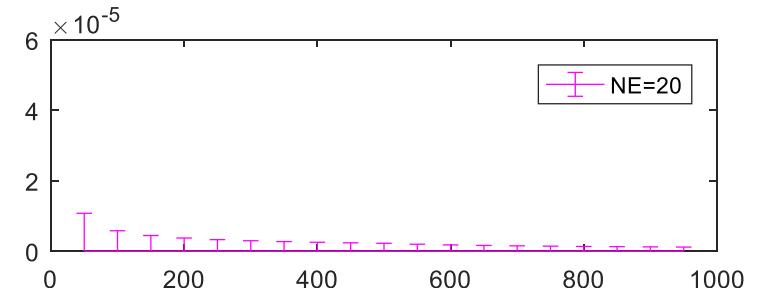
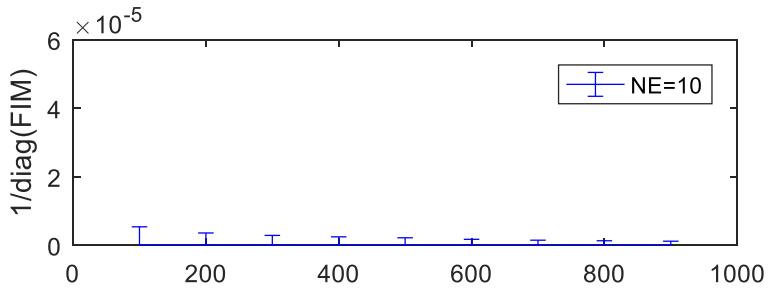
CRLB increases with increase in parameters (MOC nodes)

if $\Delta x_{\text{MOC}} < a \cdot T_c$

$$a \cdot T_c = 100 \text{ m}$$

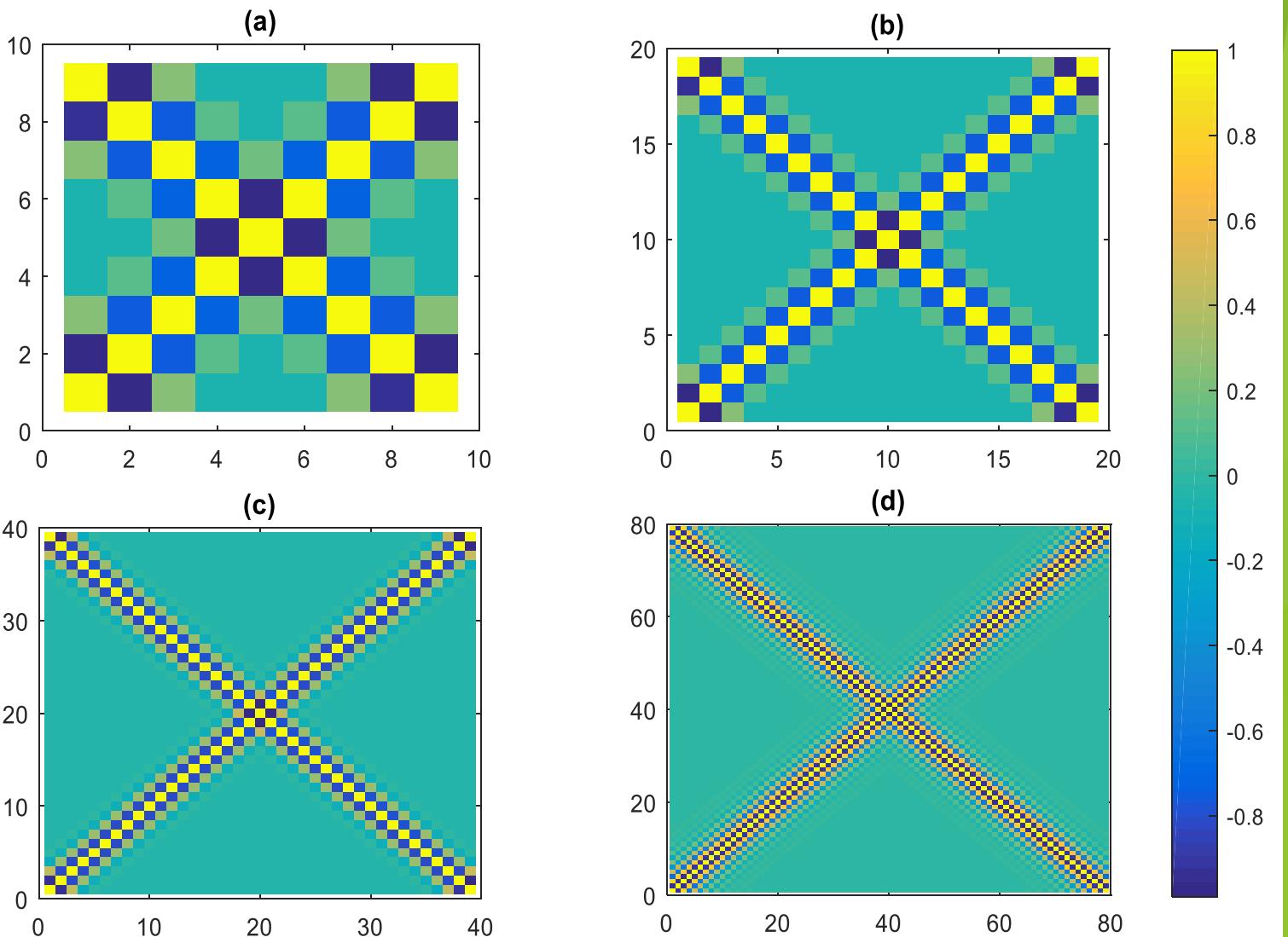


CRLB increases with increase in parameters (MOC nodes) but Fisher Information does not.

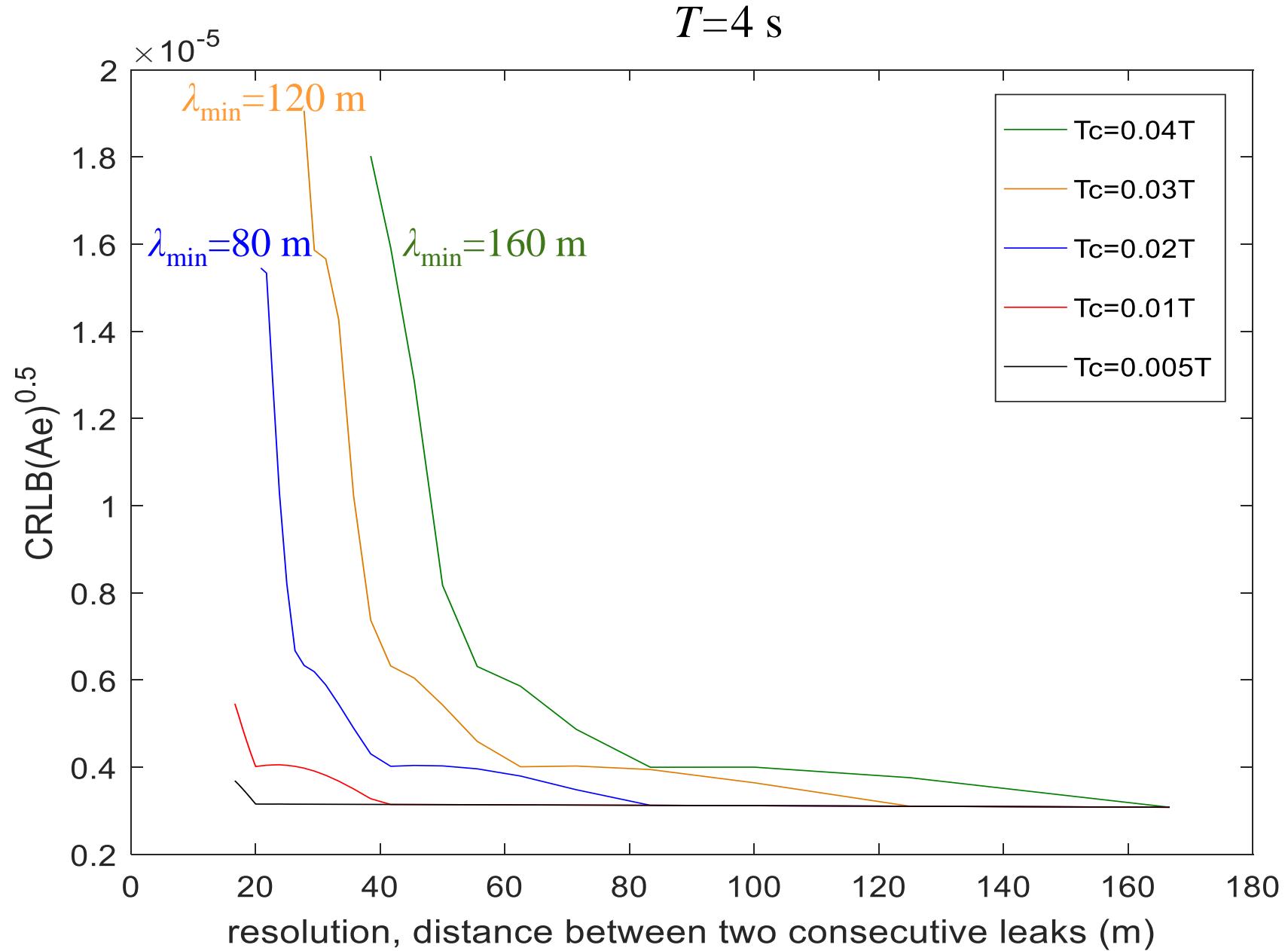


Correlation between adjacent leaks

$$r_{A_{ei}, A_{ej}} = \frac{\text{cov}(A_{ei}, A_{ej})}{\sigma_{A_{ei}} \sigma_{A_{ej}}}$$



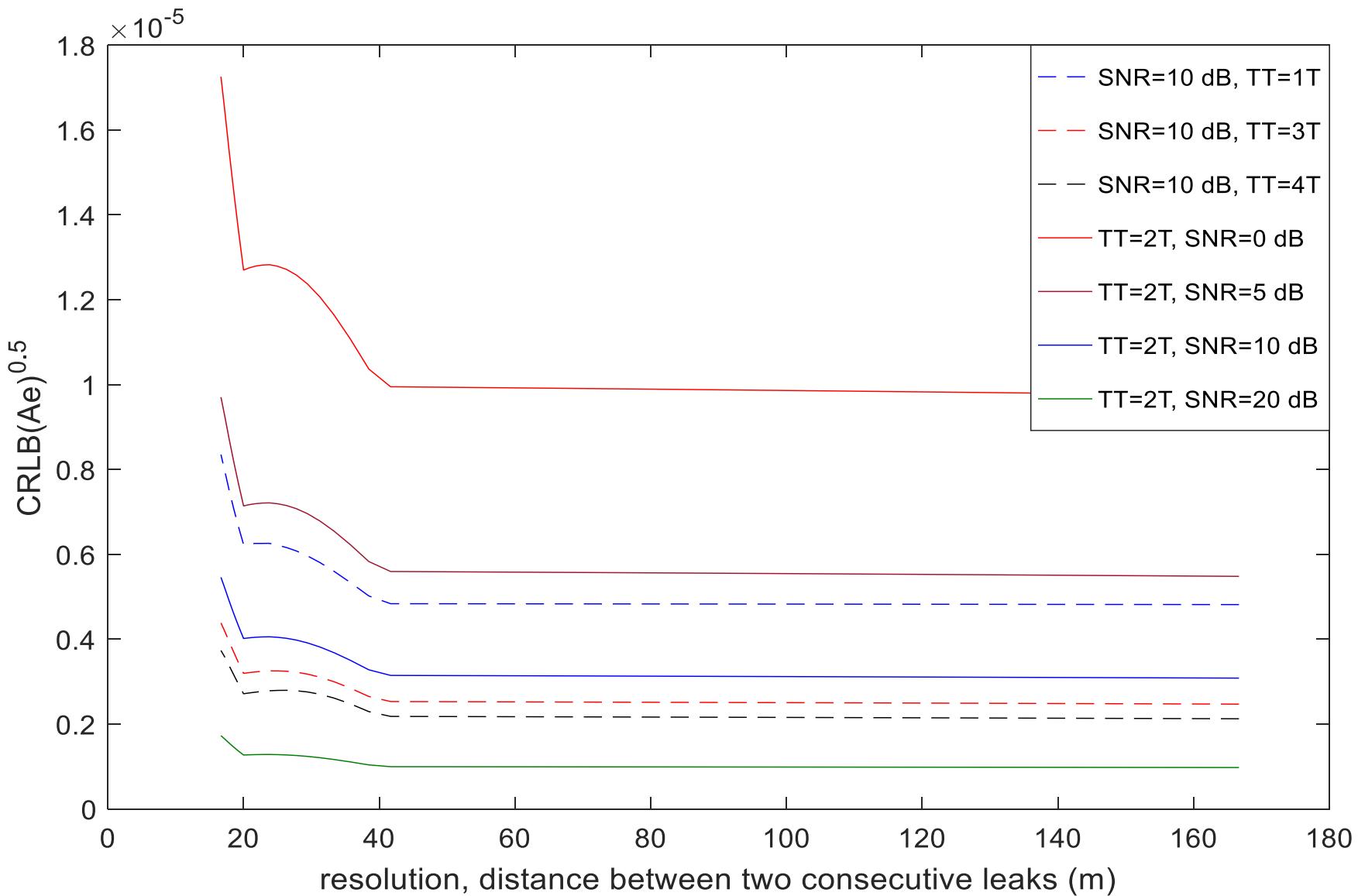
Average CRLB trend with distance between two consecutive leaks



CRLB trend with sample size and SNR

$T_c = 0.01$ $T = 0.4$ s

$a \cdot T_c = \lambda_{\min} = 40$ m



A wide-angle photograph of a park during the day. The foreground is a bright, vibrant green lawn. In the middle ground, a paved path or walkway leads through the park. On either side of the path are numerous tall, mature trees with dense green foliage. Sunlight filters through the leaves, creating bright highlights and deep shadows. The overall atmosphere is peaceful and sunny.

Thanks for your attention

Forward problem

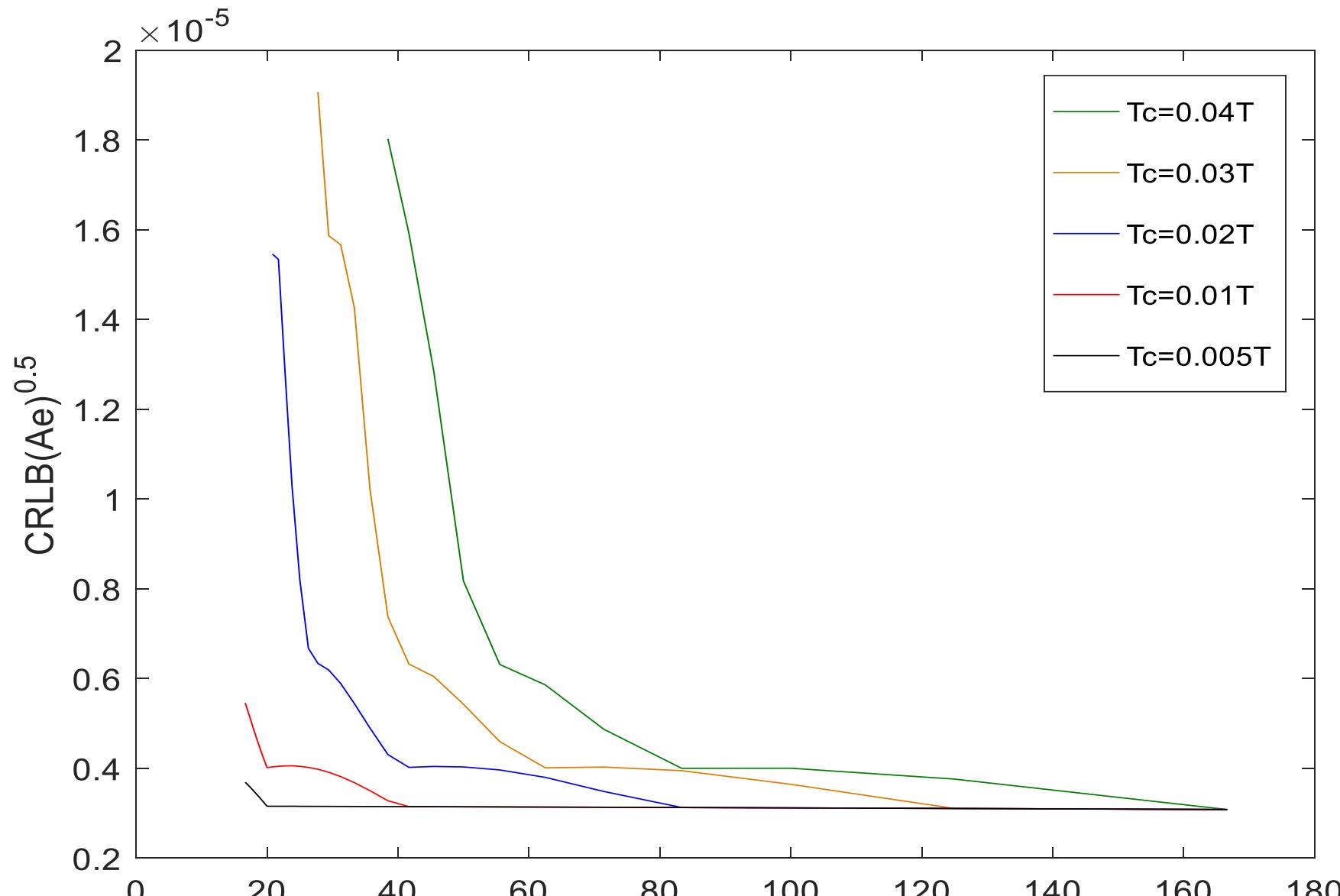
CRLB trend with distance between two consecutive leaks

CRLB increases as distance between two consecutive leaks decreases

- Determine head and flow rate (state variables)

Inverse problem

- Given governing equations and main boundary conditions
- Transient flow generation and collection of the time series data
- Optimization
- Determine some system parameters



Informally, we begin by considering an [unbiased estimator](#) $\hat{\theta}(X)$. Mathematically, "unbiased" means that

$$E[\hat{\theta}(X) - \theta | \theta] = \int (\hat{\theta}(x) - \theta) f(x; \theta) dx = 0.$$

This expression is zero independent of θ , so its partial derivative with respect to θ must also be zero. By the [product rule](#), this partial derivative is also equal to

$$0 = \frac{\partial}{\partial \theta} \int (\hat{\theta}(x) - \theta) f(x; \theta) dx = \int (\hat{\theta}(x) - \theta) \frac{\partial f}{\partial \theta} dx - \int f dx.$$

For each θ , the likelihood function is a probability density function, and therefore $\int f dx = 1$. A basic computation implies that

$$\frac{\partial f}{\partial \theta} = f \frac{\partial \log f}{\partial \theta}.$$

Using these two facts in the above lets us write

$$\int (\hat{\theta} - \theta) f \frac{\partial \log f}{\partial \theta} dx = 1.$$

Factoring the integrand gives

$$\int ((\hat{\theta} - \theta) \sqrt{f}) \left(\sqrt{f} \frac{\partial \log f}{\partial \theta} \right) dx = 1.$$

If we square the expression in the integral, the [Cauchy–Schwarz inequality](#) lets us write

$$1 = \left(\int [(\hat{\theta} - \theta) \sqrt{f}] \cdot [\sqrt{f} \frac{\partial \log f}{\partial \theta}] dx \right)^2 \leq \left[\int (\hat{\theta} - \theta)^2 f dx \right] \cdot \left[\int \left(\frac{\partial \log f}{\partial \theta} \right)^2 f dx \right].$$

The second bracketed factor is defined to be the Fisher Information, while the first bracketed factor is the expected mean-squared error of the estimator $\hat{\theta}$.

$$\text{Var}(\hat{\theta}) \geq \frac{1}{\mathcal{I}(\theta)}.$$

Fisher information matrix converge

