Rough notes on "Cramér–Rao Bounds for Three-Point Decomposition of Water and Fat" by Pineda, et al.

Jasper M. Everink

November 5, 2026, version 1

Disclaimer: These notes have not been peer-reviewed and polished. While I have used genAI to check and correct some of the results, there may still be mistakes present. Feel free to get in touch if you find an error. These notes are not written to be stand-alone to [1].

1 Intrduction

In the article [1], the authors use the Cramér-Rao bound (CRB)[1] to maximize the efficiency of the estimation of unknown water, fat, and field parameters in a three-point measurement set-up. In [1], they define an efficiency measure only to proceed with numerical optimization. These notes mathematically solve the optimization problem, verifying their found solutions whilst providing additional insight into the solutions.

2 The objective: number of signals averaged (NSA)

To measure the efficiency of estimating a parameter p_k, they define the number of signals averaged (NSA) as

$$\mathrm{NSA}(\hat{p}_k) := N \frac{\mathrm{CRB}(p_k \mid p_{l\neq k})}{\sigma^2_{\hat{p}_k}},$$

where $\mathrm{CRB}(p_k \mid p_{l\neq k})$ denotes the CRB for estimating p_k assuming the other coefficients of p are known. This quantity is closely related to the efficiency of an unbiased estimator, which for the setting where there is only a single scalar unknown is defined as the ratio between the CRB and the variance. Rigorously, let $\mathcal{I} \in \mathbb{R}^{5\times 5}$ be the Fisher information matrix (FIM) of the model with all unknowns, then simply $\mathrm{CRB}(p_k \mid p_{l\neq k}) = [\mathcal{I}_{kk}]^{-1}$. In comparison, the CRB of p_k with all unknowns satisfies $\sigma^2_{\hat{p}_k} \geq \left[\mathcal{I}^{-1}\right]_{kk}$ for any unbiased estimator satisfying basic regularity conditions. With the estimator variance $\sigma^2_{\hat{p}_k}$ being unknown, $\mathrm{NSA}(\hat{p}_k)$ is not easily computable without using Monte Carlo estimation. Therefore, we can apply the CRB to obtain an upper bound on the NSA, i.e.,

$$\mathrm{NSA}(\hat{p}_k) = N \frac{\mathrm{CRB}(p_k \mid p_{l\neq k})}{\sigma^2_{\hat{p}_k}} \leq N \frac{1}{\mathcal{I}_{kk}\left[\mathcal{I}^{-1}\right]_{kk}} =: \widetilde{\mathrm{NSA}}(\hat{p}_k).$$

This bound holds for any unbiased estimator $\hat{p}_k$ and is achieved for any efficient estimator. Numerical experiments from [1] show that this is generally not the case, but nevertheless is a good proxy for optimization. In particular, the following proposition provides the best-case upper bound and a condition for this bound to be tight.

Proposition 1. *Assume that $\mathcal{I}$ is symmetric positive definite, so that $\widetilde{\mathrm{NSA}}(\hat{p}_k)$ is well-defined. Then $\widetilde{\mathrm{NSA}}(\hat{p}_k) \leq N$, with equality if and only if $\mathcal{I}_{kj} = 0$ for all $j \neq k$.*

[1] The article [1] states "If an algorithm produces the correct estimate in the absence of noise, then in the presence of noise, the variance of that estimate cannot be lower than the CRB." The implicit implication that an algorithm that produces the correct estimate in the absence of noise is an unbiased estimator is generally not correct. See Appendix B for a theoretical setting in which this cannot hold for nonlinear problems.

Proof. Define $\Lambda_j \in \mathbb{R}^{4\times 4}$ as the sub-matrix of $\mathcal{I}$ with the j-th row and column removed and by $\lambda_j \in \mathbb{R}^4$ the j-th column of $\mathcal{I}$ with the j-th row removed. Then using the Schur complement and the non-singularity of $\mathcal{I}$, we obtain

$$\left[\mathcal{I}^{-1}\right]_{kk} = \left(\mathcal{I}_{kk} - \lambda_k^T \Lambda_k^{-1} \lambda_k\right)^{-1}, \text{ hence } \widetilde{\mathrm{NSA}}(\hat{p}_k) = N\left(1 - \frac{\lambda_k^T \Lambda_k^{-1} \lambda_k}{\mathcal{I}_{kk}}\right).$$

Because $\mathcal{I}$ is symmetric positive definite, $\mathcal{I}_{kk} > 0$ and $\lambda_k^T \Lambda_k^{-1} \lambda_k \geq 0$, hence $\widetilde{\mathrm{NSA}}(\hat{p}_k) \leq N$. Equality happens precisely when $\lambda_k = 0$, i.e., $\mathcal{I}_{kj} = 0$ for all $j \neq k$. □

Proposition 1 simplifies maximizing $\widetilde{\mathrm{NSA}}(\hat{p}_k)$ to finding the zeros of entries of the FIM, but only if at least one such solution exists. If not, then $\widetilde{\mathrm{NSA}}(\hat{p}_k) < N$ and analytical maximization becomes more difficult.

We consider the following signal model:

$$s_k = (We^{i\phi_W} + Fe^{i(\phi_F + q_k)})e^{i\psi q_k}.$$

To find zeroes, we compute the FIM at the nominal phase $\phi_W = \phi_F = 0$ in a form equivalent to [1],

$$\mathcal{I} = \begin{bmatrix} K & 0 & C & -FS & -FQ_s \\ 0 & KW^2 & WS & WFC & W(WQ + FQ_c) \\ C & WS & K & 0 & WQ_s \\ -FS & WFC & 0 & KF^2 & F(WQ_c + FQ) \\ -FQ_s & W(WQ + FQ_c) & WQ_s & F(WQ_c + FQ) & (W^2 + F^2)Q_2 + 2WFQ_{2c} \end{bmatrix},$$

where the following shorthand notation is used

$$C = \sum_{k=1}^{K} \cos(q_k), \qquad S = \sum_{k=1}^{K} \sin(q_k), \qquad Q = \sum_{k=1}^{K} q_k,$$
$$Q_c = \sum_{k=1}^{K} q_k \cos(q_k), \qquad Q_s = \sum_{k=1}^{K} q_k \sin(q_k),$$
$$Q_2 = \sum_{k=1}^{K} q_k^2, \qquad Q_{2c} = \sum_{k=1}^{K} q_k^2 \cos(q_k).$$

For $\widetilde{\mathrm{NSA}}(\hat{p}_k)$ to be well-defined, we need $\mathcal{I}$ to be non-singular. Due to Proposition 2 in the appendix, this holds when $W \neq 0$, $F \neq 0$, $W \neq F$ and q_k are distinct modulo 2π. Before considering the separate cases of magnitude, phase and field, we can scale the entries of $\mathcal{I}$ such that its zeros are easier to compute. So consider

$$\tilde{\mathcal{I}} = \begin{bmatrix} 1 & 0 & C & S & Q_s \\ 0 & 1 & S & C & WQ + FQ_c \\ C & S & 1 & 0 & Q_s \\ S & C & 0 & 1 & WQ_c + FQ \\ Q_s & WQ + FQ_c & Q_s & WQ_c + FQ & (W^2 + F^2)Q_2 + 2WFQ_{2c} \end{bmatrix}.$$

The first four columns, without the diagonal entries, all yield the joint conditions $C = 0$ and $S = 0$. Combining these results in

$$0 = C + iS = \sum_{j=1}^{K} \cos(q_j) + i\sin(q_j) = \sum_{j=1}^{K} \exp(iq_j).$$

For the case $K = 3$, the above equality holds precisely when $q_k = c + k\frac{2\pi}{K} + m_k 2\pi$, for any $c \in \mathbb{R}$ and $m_k \in \mathbb{Z}$. Though we choose $m_k = 0$ for all k such that the acquisition range is small. Thus, for optimal water/fat magnitude and phase with $K = 3$, e^{iq_k} must be equally spaced around the unit circle. For $K \geq 4$, this uniformity condition does not need to hold, but is sufficient, hence we will focus on the uniformly spaced points when relevant.

Water and fat magnitudes

The off-diagonal column vectors $\tilde{\lambda_1}$ and $\tilde{\lambda_3}$ of $\tilde{\mathcal{I}}$, which correspond to water and fat magnitudes respectively, have the same zeros. The conditions $C = 0$ and $S = 0$, which correspond to the cross-terms between water and fat, are satisfied when the q_k to be uniformly distributed around the circle. The precise offset c can be derived from the final condition $Q_s = 0$. In particular, assuming $m_k = 0$ and the uniformly distributed points, both Q_c and Q_s simplify to

$$Q_c = \pi \csc\left(\frac{\pi}{K}\right) \cos\left(c - \frac{\pi}{2} + \frac{\pi}{K}\right), \text{ and } Q_s = \pi \csc\left(\frac{\pi}{K}\right) \sin\left(c - \frac{\pi}{2} + \frac{\pi}{K}\right). \tag{1}$$

Hence $Q_s = 0$, results in $c = l\pi + \frac{\pi}{2} - \frac{\pi}{K}$ for any $l \in \mathbb{Z}$. For $K = 3$ this exactly recovers the solutions from [1],

$$\begin{bmatrix} q_1 & q_2 & q_3 \end{bmatrix} = \begin{bmatrix} -\frac{\pi}{6} + l\pi & \frac{\pi}{2} + l\pi & \frac{7\pi}{6} + l\pi \end{bmatrix},$$

and uniqueness assuming $m_k = 0$. This shows that the solutions found in [1] are the unique optima for $K = 3$. In general, q_k are those uniformly distributed modulo 2π such that at least one q_k is of the form $l\pi + \frac{\pi}{2}$ are optimal, i.e., the relative phases between water and fat are orthogonal, which aligns with the intuition used in practice.

Water and fat phases

Similarly, for water and fat phase, it is sufficient to consider uniformly distributed points on the circle with an unknown offset $c \in \mathbb{R}$. This offset is given by $WQ + FQ_c = 0$ for water phase and $WQ_c + FQ = 0$ for fat phase. To obtain a result independent of the water and fat magnitudes, we want to find c such that both $Q = 0$ and $Q_c = 0$, which yields zeros for both water and fat phase simultaneously. From previously discussed arguments $Q_c = 0$ results in $c = l\pi - \frac{\pi}{K}$ for any $l \in \mathbb{Z}$. For $Q = 0$, we get $l = 1$. This recovers the solution from [1] as a solution,

$$\begin{bmatrix} q_1 & q_2 & q_3 \end{bmatrix} = \begin{bmatrix} -\frac{2\pi}{3} & 0 & \frac{2\pi}{3} \end{bmatrix}.$$

Field parameter

For the field parameter we would need $Q_s = 0$, $WQ + FQ_c = 0$ and $WQ_c + FQ = 0$, or for conditions independent of W and F, we need $Q_s = 0$, $Q_c = 0$ and $Q = 0$. This time, we do not have a uniformity assumption, and even stronger, Q_s and Q_c cannot be both zero for a uniform solution due to (1). Combining $Q_s = 0$ and $Q_c = 0$ results in

$$\sum_{k=1}^{K} q_k \exp(iq_k) = 0.$$

For simplicity, consider $K = 3$. First assume that $q_2 = 0$, then $Q = 0$ implies $q_1 = -q_3$. The second condition therefore becomes $q_1 \exp(iq_1) = q_1 \exp(-iq_1)$, which holds when $q_1 = l\pi$ for $l \in \mathbb{Z}$. Thus, all these symmetric three-point solutions are

$$\begin{bmatrix} q_1 & q_2 & q_3 \end{bmatrix} = \begin{bmatrix} -l\pi & 0 & l\pi \end{bmatrix},$$

and these solutions contain the triple $(-\pi, 0, \pi)$ from [1]. Direct calculations show that for these solutions, the FIM is nonsingular when l is odd, hence the CRB is valid.

If all q_k are non-zero, then $Q = 0$ forces at least one of the q_k to be negative. Without loss of generality assume that $q_3 < 0$ and $q_1, q_2 > 0$. The equation $Q = 0$ implies that $q_3 = -(q_1 + q_2)$, such that $Q_s = 0$ and $Q_c = 0$ are equivalent to $q_1 e^{iq_1} + q_2 e^{iq_2} = (q_1 + q_2) e^{-i(q_1+q_2)}$. Alternatively, we can write this as

$$\frac{q_1}{q_1 + q_2} e^{iq_1} + \frac{q_2}{q_1 + q_2} e^{iq_2} = e^{-i(q_1+q_2)}.$$

This implies that $e^{-i(q_1+q_2)}$ is a strict convex combination of e^{iq_1} and e^{iq_2}, all three are points on the unit circle. Hence this can only happen if all three points are equal, i.e., $e^{iq_1} = e^{iq_2} = e^{-i(q_1+q_2)}$. Thus $e^{iq_1} = e^{iq_2} = e^{iq_3}$ which holds precisely when $q_k = c + 2\pi n_k$ for any $n_k \in \mathbb{Z}$. However, $\det(\mathcal{I}) = 0$ for these solutions, hence the CRB is not valid. Thus the symmetric three-point solutions are the only valid ones.

References

[1] A. R. Pineda, S. B. Reeder, Z. Wen, and N. J. Pelc. Cramér–rao bounds for three-point decomposition of water and fat. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 54(3):625–635, 2005.

A Non-singularity of the Fisher information matrix

The bound $\widetilde{\text{NSA}}$ is only valid when the Fisher information matrix is non-singular. This can result in various edge cases, e.g., [1] states "We also added a perturbation of 10^{-6} to the water density to avoid the numerical instability when $W = F$." To be exact, the Fisher information matrix $\mathcal{I}$ is singular when $W = F$ for some of the optimal q_k, such as for the water and fat phase. The following result summarizes some basic, non-exhaustive conditions that guarantee the non-singularity of the Fisher information matrix.

Proposition 2. *If $W > 0$, $F > 0$, $W \neq F$, $K = 3$ and the q_k are distinct modulo 2π, then the Fisher information matrix $\mathcal{I}$ is non-singular.*

Proof. The Fisher information matrix can be written as $\mathcal{I} = \text{Re}(\mathcal{J}^\star \mathcal{J})$, where $\mathcal{J} \in \mathbb{C}^{K\times 5}$ has rows

$$\mathcal{J}_{k:} = \begin{bmatrix} 1 & iW & e^{iq_k} & iFe^{iq_k} & iq_k(W + Fe^{iq_k}) \end{bmatrix}.$$

Using this decomposition, $\mathcal{I}$ is singular precisely when there exists a non-zero vector $x \in \mathbb{R}^5$ such that $\mathcal{J}x = 0$, i.e.,

$$x_1 + x_2 iW + x_3 e^{iq_k} + x_4 iFe^{iq_k} + x_5 iq_k(W + Fe^{iq_k}) = 0, \text{ for } k = 1, \dots, K.$$

Take any $x \in \mathbb{R}^5$ such that $\mathcal{J}x = 0$. We proceed by considering two cases. Assume $x_5 = 0$, then $(x_1 + x_2 iW) + (x_3 + x_4 iF)e^{iq_k} = 0$ for $k = 1, \dots K$. Writing $\alpha = x_1 + x_2 iW$, $\beta = x_3 + x_4 iF$, and $z_k = e^{iq_k}$, this equals $\alpha + \beta z_k = 0$ for $k = 1, \dots K$. As a non-zero first-degree polynomial $z \mapsto \alpha + \beta z$ can only have at most one zero, and because we assume that $K = 3$, and the z_k are distinct, we must have that $\alpha = \beta = 0$. Hence $x_1 = x_2 = x_3 = x_4 = 0$, and $x = 0$.

Alternatively, assume that $x_5 \neq 0$, rescaling the other coefficients by ix_5, and using the previous α, β and q_k conventions, we get

$$\alpha + \beta z_k + q_k(W + Fz_k) = 0, \text{ for } k = 1, \dots, K,$$

or equivalently $WA + FB = 0$, where

$$A = \det \begin{bmatrix} 1 & z_1 & q_1 \\ 1 & z_2 & q_2 \\ 1 & z_3 & q_3 \end{bmatrix}, \text{ and } B = \det \begin{bmatrix} 1 & z_1 & q_1 z_1 \\ 1 & z_2 & q_2 z_2 \\ 1 & z_3 & q_3 z_3 \end{bmatrix}.$$

Using that $B = -\bar{A} \prod_{k=1}^K z_k$, results in the $WA = F \prod_{k=1}^K z_k \bar{A}$. But in absolute value, this requires $|W||A| = |F||A|$.

If $A = 0$, then $q_k = \gamma + \delta z_k$ for $k = 1, \dots, K$. If $\delta = 0$, then $q_k = \gamma$, which is a contradiction. Alternatively, if $\delta \neq 0$, then all z_k must lie on a single line that intersects the unit circle, which cannot happen if $K \geq 3$ and the z_k are distinct. Hence $|A| \neq 0$.

Then $|W| = |F|$, but this is a contradiction by assumption. Hence we can conclude that $x = 0$, and $\mathcal{I}$ is non-singular. □

B Unbiased estimator

The article [1] states: "If an algorithm produces the correct estimate in the absence of noise, then in the presence of noise, the variance of that estimate cannot be lower than the CRB." As the CRB requires an unbiased estimator, this statement implicitly assumes that an algorithm that computes the correct parameters from uncorrupted data is unbiased. Denoting this algorithm by θ, this assumes $\theta(A(x)) = x$, i.e., θ is the left inverse of A. If this holds for all $x \in \mathbb{R}^n$ and A is

assumed to be surjective, then this already forces the strong conclusion $\theta = A^{-1}$. If we furthermore assume the estimator is unbiased, then this severely limits the possible functions θ. The following proposition shows that in the setting above, if θ grows at most linearly, then θ must be affine. This also forces A to be affine, thus such an algorithm cannot exist for nonlinear, surjective problems.

Proposition 3. *Assume $A : \mathbb{R}^n \to \mathbb{R}^m$ is surjective and $\Sigma \in \mathbb{R}^{m\times m}$ symmetric positive definite. Let $y = A(x)$, $\bar{y} \sim \mathcal{N}(y, \Sigma)$, and θ be measurable such that for all $x \in \mathbb{R}^n$, $\theta(A(x)) = x$, $\mathbb{E}[\theta(\bar{y})] = x$, and θ grows at most linearly, i.e., $\|\theta(y)\|_2 \leq C(1 + \|y\|_2)$ for some $C > 0$. Then, θ must be affine.*

Proof outline. By surjectivity of A and assumptions on θ, we have that $\theta(y) = \mathbb{E}[\theta(y + e)]$ with $e \sim \mathcal{N}(0, \Sigma)$. Equivalently, we can write $\theta = \theta \star \psi_\Sigma$, where $\star$ is the convolution operator and ψ_Σ is the density of a zero-mean multivariate Gaussian with covariance matrix Σ. Taking the Fourier transform (in the sense of tempered distributions) results in $\hat{\theta} = \hat{\theta} \cdot \widehat{\psi_\Sigma}$, equivalently, $(1 - \widehat{\psi_\Sigma}) \cdot \hat{\theta} = 0$. Because $(1 - \widehat{\psi_\Sigma})(\xi) \neq 0$ whenever $\xi \neq 0$, we must have $\operatorname{supp} \hat{\theta} \subseteq \{0\}$. Such $\hat{\theta}$ is necessarily a finite linear combination of Dirac delta distributions and their derivatives, hence θ must be a polynomial. By assumption, the growth of θ is at most linear, hence θ must be an affine map. □