1. Physics of imaging
Unconventional imaging and co-design
Part IV: Cramer Rao Lower Bound in coherent imaging

Corinne Fournier

Equipe ODIR (Optical Design and Image Reconstruction),
Laboratoire Hubert Curien, UMR CNRS 5516
corinne.fournier@univ-st-etienne.fr

june 2021

Outline

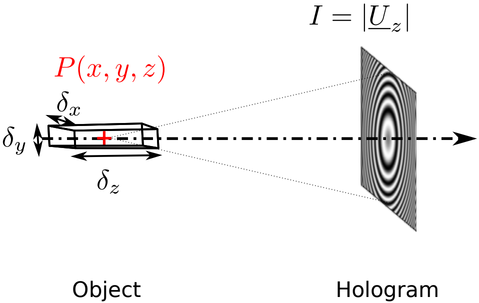
1. Cramer Rao Lower Bound in coherent imaging
   - Context / Issue
   - Statistical approach to accuracy estimation
   - Example in Digital Holography
   - Conclusion

Cramer Rao Lower Bound in coherent imaging

Context / Issue

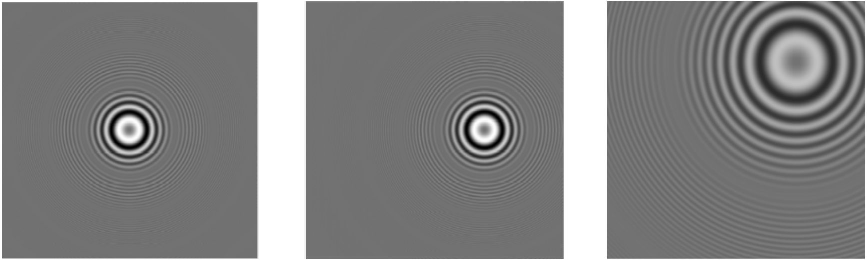
Positioning accuracy

In digital holography, imaging system can be considered has a optical/numerical system.



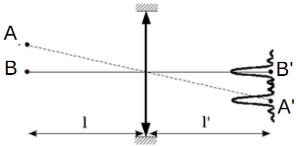
Positioning accuracy

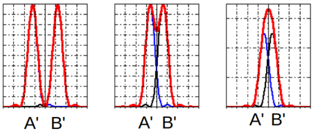
In digital holography, imaging system can be considered has a optical/numerical system.
The positioning accuracy depends on the position of the object : z,x and y



Two points resolution

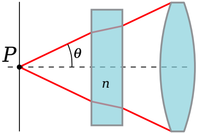
Two-points resolution measures
the capability of the system to separate two points sources e.g : Rayleigh criterion or Sparrow criterions





$$\Delta_x \propto \frac{\lambda}{\Omega}, \quad \Delta_y \propto \frac{\lambda}{\Omega}, \quad \Delta_z \propto \frac{\lambda}{\Omega^2}$$

where λ is the wavelength of the light and $\Omega = sin(\theta)$ is the Numerical Aperture of the imaging system



Two points resolution

The positioning accuracy depends also on the imaging system PSF. It is also called single point resolution.
The Rayleigh criterion or Sparrow criterion are empirical and do not account for SNR.

Statistical approach to accuracy estimation

Maximum Likelihood Estimation (MLE)

The likelihood in the case of additive White Gaussian Noise (with variance σ^2) is :

$$p(\mathbf{m};\theta) \propto exp\left(-\frac{1}{2\sigma^2}[\mathbf{d}-\mathbf{m}_\theta]^t[\mathbf{d}-\mathbf{m}_\theta]\right)$$

where :

- $\mathbf{d}$ is the data vector of N components.
- $\mathbf{m}$ is the model vector depending on θ (vector of parameters, e.g. $\theta = (x, y, z)^t$).

The ML estimate θ^* is the vector that maximizes the likelihood.

$\mathbf{m}_{\theta^*}$ is the model that best fits the data.

Asymptotic behaviour of MLE

- Unbiased (i.e., bias $\underset{N \to +\infty}{\longrightarrow} 0$)
- Efficient (i.e., reaches Cramér-Rao lower bound when $N \to +\infty$)
- Normally distributed : $\hat{\theta}^{ML} \underset{N \to +\infty}{\longrightarrow} N(\theta^*, \mathbf{I}^{-1}(\theta^*))$

* S. M. Kay, "Fundamentals of Statistical Signal Processing : Estimation Theory", Prentice-Hall, 2005.

Cramér-Rao Lower Bound (CRLB)

The variance of $\hat{\theta}$ is lower bounded by the inverse of the Fisher information matrix $I(\theta^*)$:

$$var(\theta_i) \geqslant [\mathbf{I}^{-1}(\boldsymbol{\theta}^*)]_{i,i}$$

with Fisher information matrix :

$$[\mathbf{I}(\boldsymbol{\theta})]_{i,j} \equiv -E\left[\frac{\partial^2 \ln p(\mathbf{m};\boldsymbol{\theta})}{\partial\theta_i \partial\theta_j}\right]$$

wich is connected to the curvature of the log-likelihood function.

* P. Réfrégier, "Noise Theory and Application to Physics : From Fluctuations to Information", Springer Verlag, 2004.

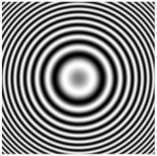
Example in Digital Holography

Single point resolution on the optical axis

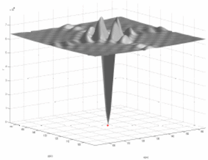
The diffraction pattern of a point source located at (x, y, z) is :

$$m(x, y, z) = \alpha.sin\left(\pi\frac{(x_k - x)^2 + (y_k - y)^2}{\lambda z}\right)$$

with (x_k, y_k) the coordinate of the observation point.
On the optical axis $x = y = 0$:



Hologram



Log_likelihood function

* C. Fournier, L. Denis, T. Fournel, "On the single point resolution of on-axis digital holography." JOSA A, 2010.

Single point resolution on the optical axis

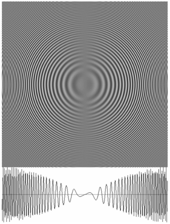
$$[\mathbf{I}(\theta)]_{i,j} \equiv \frac{1}{2\sigma^2} E\left[\frac{\partial^2}{\partial\theta_i\partial\theta_j}(\mathbf{d}-\mathbf{m}_\theta)^t(\mathbf{d}-\mathbf{m}_\theta)\right] = \frac{1}{\sigma^2}\left(\frac{\partial \mathbf{m}_\theta}{\partial\theta_i}\right)^t\left(\frac{\partial \mathbf{m}_\theta}{\partial\theta_i}\right)$$

Ignoring the sampling on the sensor, $\mathbf{I}$ has an analytical form :

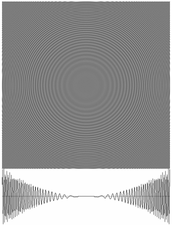
$$[\mathbf{I}(\theta)]_{i,j} \equiv \frac{1}{\sigma^2 . L^2} \iint_{-L/2,-L/2}^{L/2,L/2} \left(\frac{\partial m_{x,y,z}(x_k,y_k)}{\partial\theta_i} . \frac{\partial m_{x,y,z}(x_k,y_k)}{\partial\theta_j}\right)\Bigg|_{x=0}^{y=0} dx_k\, dy_k$$

Illustration :

$\left(\frac{\partial \mathbf{m}}{\partial x}\right)$



$\left(\frac{\partial \mathbf{m}}{\partial z}\right)$



Single point resolution on the optical axis

Hypothesis : $L^2 \gg \lambda z$.

The analytical form of Fisher matrix is :

$$\mathbf{I}(x,y,z) = \frac{\alpha^2}{\sigma^2}\begin{pmatrix} \frac{\pi^2 L^2}{6\lambda^2 z^2} & 0 & 0 \\ 0 & \frac{\pi^2 L^2}{6\lambda^2 z^2} & 0 \\ 0 & 0 & \frac{7\pi^2 L^4}{360\lambda^2 z^4} \end{pmatrix}$$

Fisher information matrix is diagonal

→the covariance matrix is also diagonal

(estimation errors on the parameters x, y and z are not correlated)

Single point resolution on the optical axis

Covariance matrix : $\begin{pmatrix} \sigma_x^2 & 0 & 0 \\ 0 & \sigma_y^2 & 0 \\ 0 & 0 & \sigma_z^2 \end{pmatrix}$

$\sigma_x = \sigma_y = c \frac{\sigma}{\alpha} \frac{\lambda}{L/z}$ and $\sigma_z = 6\sqrt{5/21}\, c \frac{\sigma}{\alpha} \frac{\lambda}{(L/z)^2}$

with $c = \sqrt{6}/\pi$

→ This result is in agreement with the classical formulas :

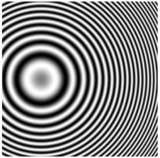
→ The resolution improves proportionally with the signal to noise ratio : $\frac{\alpha}{\sigma}$

Single point resolution out of the optical axis

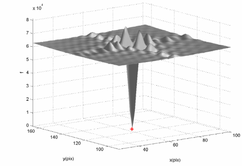
The diffraction pattern of a point source located at (x, y, z) is :

$$m(x, y, z) = \alpha.sin\left(\pi\frac{(x_k - x)^2 + (y_k - y)^2}{\lambda z}\right)$$

with (x_k, y_k) the coordinate of the observation point.
Out of the optical axis $x \neq 0, y \neq 0$:



Hologram
Model



Log-likelihood function

Analytical form of the Fisher information matrix

$$[\mathbf{I}(\theta)]_{i,j} \equiv \frac{1}{\sigma^2.L^2} \iint_{-L/2,-L/2}^{L/2,L/2} \left(\frac{\partial m_{x,y,z}(x_k, y_k)}{\partial \theta_i} . \frac{\partial m_{x,y,z}(x_k, y_k)}{\partial \theta_j} \right) \Bigg|_{x=0}^{y=0} dx_k dy_k$$

Under the hypothesis $L^2 \gg \lambda z$, the analytical form of Fisher matrix is :

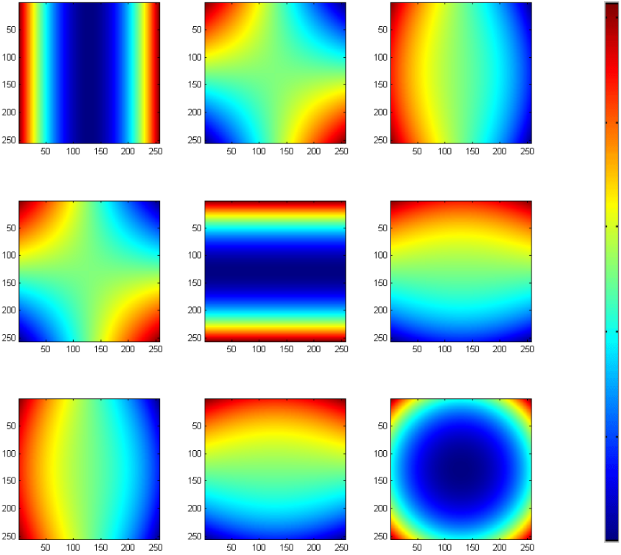
$$\frac{\alpha^2}{\sigma^2} \begin{pmatrix} \frac{\pi^2(L^2+12x^2)}{6\lambda^2 z^2} & \frac{2\pi^2 xy}{\lambda^2 z^2} & -\frac{\pi^2 x(L^2/3+x^2+y^2)}{\lambda^2 z^3} \\ \frac{2\pi^2 xy}{\lambda^2 z^2} & \frac{(\pi^2 L^2+12y^2)}{6\lambda^2 z^2} & -\frac{\pi^2 y(L^2/3+x^2+y^2)}{\lambda^2 z^3} \\ -\frac{\pi^2 x(L^2/3+x^2+y^2)}{\lambda^2 z^3} & -\frac{\pi^2 y(L^2/3+x^2+y^2)}{\lambda^2 z^3} & \frac{7\pi^2\left(L^4+120L^2(x^2+y^2)/7+180(x^2+y^2)^2/7\right)}{360\lambda^2 z^4} \end{pmatrix}$$

Fisher information matrix is no more diagonal

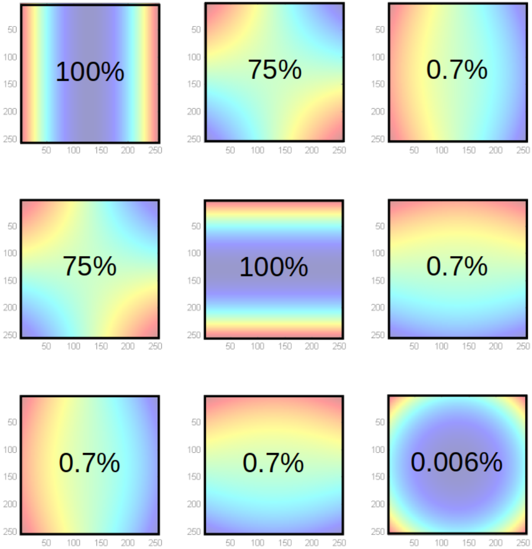
→the covariance matrix is also not diagonal

(estimation errors on the parameters x, y and z are correlated !)

Fisher information matrix (x,y) maps



Fisher information matrix (x,y) maps



Analytical form of the covariance matrix

$$\begin{pmatrix} \left(1+156\frac{x^2L^2}{K}\right)\sigma_x^2 & 156\frac{x\,y\,L^2}{K}\sigma_x\sigma_y & \sqrt{\frac{7}{15}}L\,x\,\frac{60L^2+180x^2+180y^2}{K}\,\sigma_x\,\sigma_z \\ 156\frac{x\,y\,L^2}{K}\sigma_x\sigma_y & \left(1+156\frac{y^2L^2}{K}\right)\sigma_y^2 & \sqrt{\frac{7}{15}}L\,y\,\frac{60L^2+180x^2+180y^2}{K}\,\sigma_y\,\sigma_z \\ \sqrt{\frac{7}{15}}L\,x\,\frac{60L^2+180x^2+180y^2}{K}\,\sigma_x\,\sigma_z & \sqrt{\frac{7}{15}}L\,y\,\frac{60L^2+180x^2+180y^2}{K}\,\sigma_y\,\sigma_z & \left(1+\frac{120(x^2+y^2)L^2-180(x^2+y^2)^2}{K}\right)\sigma_z^2 \end{pmatrix}$$

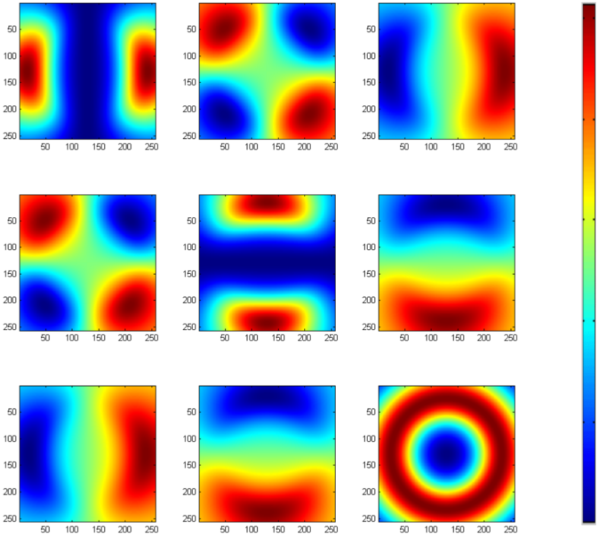
$$K = 7L^4 + \left(-36x^2 - 36y^2\right)L^2 + 180(x^2+y^2)^2$$

→ This result is no more in agreement with the classical formulas of on-axis resolution:

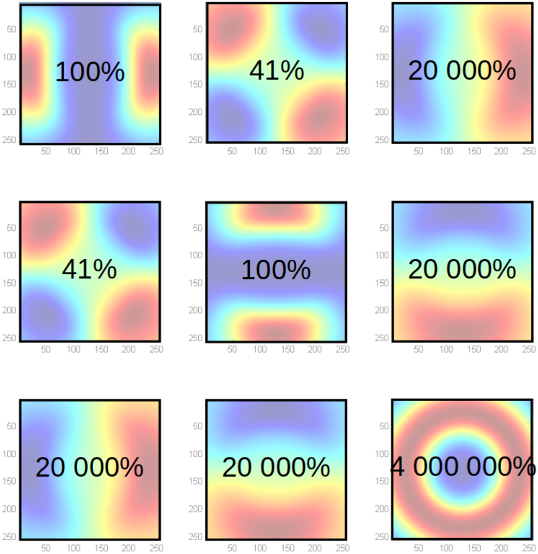
$\cancel{\sigma_x \propto \frac{\lambda}{\Omega}}$ and $\cancel{\sigma_z \propto \frac{\lambda}{\Omega^2}}$

→ The resolution still improves proportionally with the signal to noise ratio : $\dfrac{\alpha}{\sigma}$

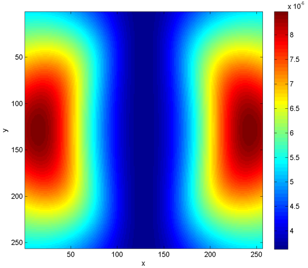
Covariance matrix (x,y) maps

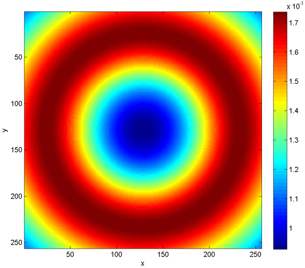


Covariance matrix (x,y) maps



Standard deviation maps





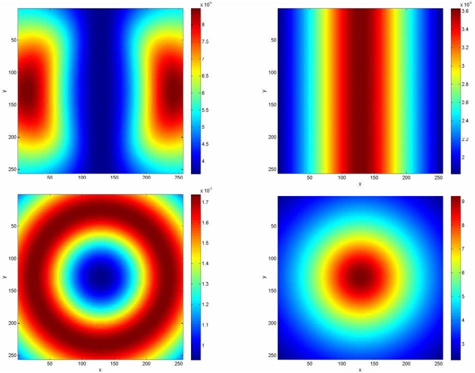
Standard deviation on x and z

Comparison with "intuitive" results (no correlation between parameters $x$and z – Fisher matrix is diagonal)

Root Mean square error of x

$\sqrt{V_{xx}}$

$1/\sqrt{I_{xx}}$



Root Mean square error of z

$\sqrt{V_{zz}}$

$1/\sqrt{I_{zz}}$

Conclusion

1. Cramer Rao Lower Bound in coherent imaging
   - Context / Issue
   - Statistical approach to accuracy estimation
   - Example in Digital Holography
   - Conclusion

Conclusion

+

- Closed form expression of lower bounds in inline holography (using a simple model), agreement with Rayleigh criterion on the optical axis.
- Correlation between parameters (out of the optical axis) not intuitive
- Same methodology can be applied for more complex parametric model (e.g., accounting for pixel integration, spherical object, beads in astigmatic setup, ...)
- Useful for a better conditioning of the Parametric IP or stopping criterion in parameters optimization
- Performance model for codesign (bien sûr !).

–

- Global minimum is considered achievable.
- Noise was considered additive, white and Gaussian.
- Numerical gradients not always easy to compute.

References

- C. R. Rao, "Information and the accuracy attainable in the estimation of statistical parameters." Reson. J. Sci. Educ 20, 1945.
- S. M. Kay, "Fundamentals of Statistical Signal Processing : Estimation Theory", Prentice-Hall, 2005.
- P. Réfrégier, "Noise Theory and Application to Physics : From Fluctuations to Information", Springer Verlag, 2004.
- J.A. Fessler, A.O. Hero, "Cramér-Rao lower bounds for biased image reconstruction.", In Proceedings of 36^{th} Midwest Symposium on Circuits and Systems, pages 253-256. IEEE, 1993.
- K. Todros, J. Tabrikian, "General Classes of Performance Lower Bounds for Parameter Estimation-Part I : Non-Bayesian Bounds for Unbiased Estimators." IEEE Trans. Inform. Theory, 56(10), 2010.
- K. Todros, J. Tabrikian, "General Classes of Performance Lower Bounds for Parameter Estimation-Part II : Bayesian Bounds." IEEE Trans. Inform. Theory, 56(10), 2010.