Information and Decision Theory applied to Astronomy

Astronomía
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS UNIVERSIDAD DE CHILE

MILLENNIUM INSTITUTE OF ASTROPHYSICS

Astrometry & Photometry

Rene Mendez – UChile/ESO, Jorge Silva, Rodrigo Lobos – UChile

rmendez@u.uchile.cl

Parameter estimation & Information and Decision Theory

- Optimal parameter estimation can, in general, be formulated using "decision-making" theory (Cover & Thomas 2006),
- This framework can also provide absolute lower bounds to the uncertainty of these estimators,
- The Cramer-Rao (CR) lower uncertainty bound determines the minimum theoretical variance achievable by *any* unbiased estimator (Stuart, Ord, & Arnold 2004),
- The CR bound does not specify, *a priori*, how to construct estimators that reach that bound,
- However, Bayesian statistics allows us to write down parameter estimation expressions that *are guaranteed* to achieve a minim mean-square error (the conditional expectation, Gray and Davisson 2005),
- By contrast, classical estimation techniques using least-squares (LS) or maximum-likelihood *do not achieve*, in general, neither the CR bound (Mendez *et al.* 2013) nor the minimum Bayesian MSE, and are thus less efficient statistically. One can theoretically demonstrate however that the CR bound is *nearly reached* in many instances using a LS estimator (see talk by graduate student Rodrigo Lobos on Tuesday afternoon),
- We are working on implementing optimal Bayesian estimators for photometry and astrometry (see talk by student Alex Echeverria on Tuesday afternoon).

Our contribution, so far

- The IDS group (see box on the right, R. Mendez, J. Silva & DIE & DAS students R. Lobos & R. Oróstica) have been collaborating for two years on the application of the CR bound to the problem of astrometry and photometry from CCD imaging data,
- Results have been published in the PASP Journal by Mendez, *et al.* 2013 & 2014,
- For mathematical simplicity, we have so far analyzed the case of a linear-detector in a 2-dimensional (2d) estimation scenario: Astrometric location and total flux of the source,
- Our CR estimates agree very well with the empirically determined maximum precisions achieved in photometry and astrometry: Astrometric accuracy *vs.* flux, ground (HST) & space-based optical astrometry, photometric precision at extremely high S/N,
- Thus, the CR can be used as a benchmark to guide in: Detector design, testing of pipelines, science planning and design of observational strategies,
- We plan to extend these studies to incorporate more subtle observational features such as non-uniform pixel response functions, non-symmetric PSFs, rapidly fluctuating backgrounds, moving targets, radiation damage and CTE effects, among others.

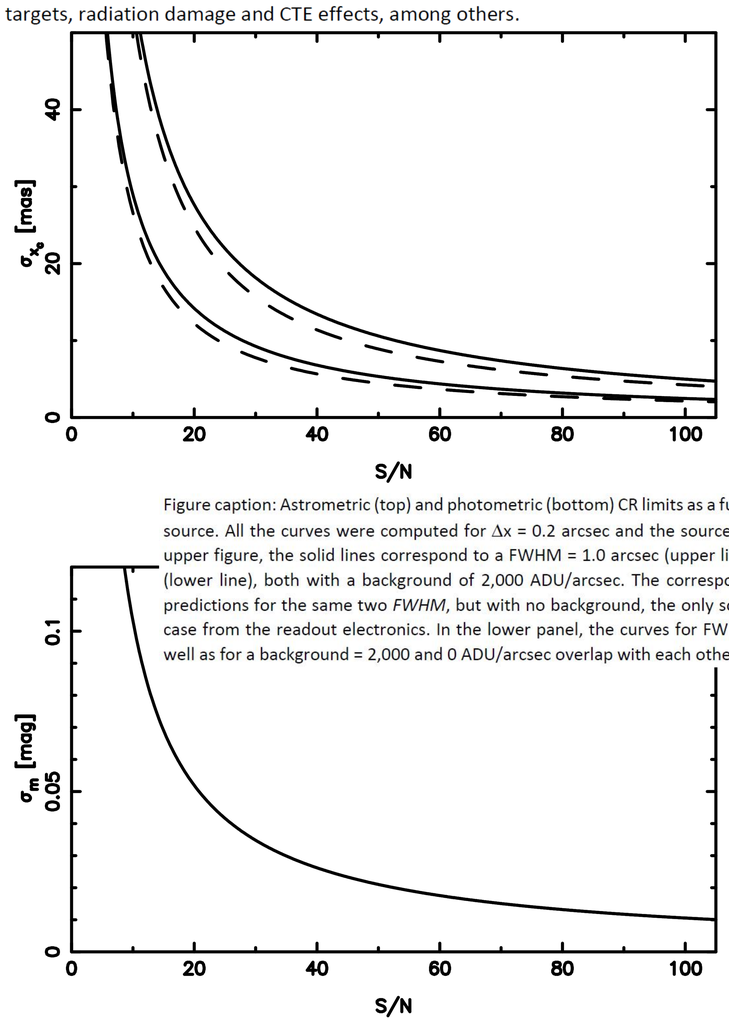


Figure caption: Astrometric (top) and photometric (bottom) CR limits as a function of *S/N* for a Gaussian source. All the curves were computed for Δx = 0.2 arcsec and the source centered on a pixel. In the upper figure, the solid lines correspond to a FWHM = 1.0 arcsec (upper line) and FWHM = 0.5 arcsec (lower line), both with a background of 2,000 ADU/arcsec. The corresponding dashed lines are the predictions for the same two *FWHM*, but with no background, the only source of noise comes in this case from the readout electronics. In the lower panel, the curves for FWHM = 1.0 and 0.5 arcsec, as well as for a background = 2,000 and 0 ADU/arcsec overlap with each other.

Application of the CR bound to the Photometry + Astrometry problem on CCD detectors - II

(Mendez *et al.* 2013, 2014)

In the high-resolution regime, the CR bound can be *approximated* by:

- For flux-dominated sources (F/B >>1):

$$\sigma^2_{\tilde{F}} = [\mathcal{I}_{(x_c,\tilde{F})}(n)^{-1}]_{2,2} \approx GF \cdot \left(1 + n \cdot \frac{B}{F}\right)$$

$$\sigma^2_{x_c} = [\mathcal{I}_{(x_c,\tilde{F})}(n)^{-1}]_{1,1} \approx \frac{1}{8\ln 2} \cdot \frac{1}{GF} \cdot \left(1 + \frac{1}{8\ln 2} \cdot \frac{B}{F} \cdot \frac{\sum_{i=1}^{n}(x_i - x_c)^2}{FWHM^2}\right) \cdot FWHM^2$$

- For background-dominated sources (F/B<<1):

$$\sigma^2_{\tilde{F}} = [\mathcal{I}_{(x_c,\tilde{F})}(n)^{-1}]_{2,2} \approx \sqrt{\frac{\pi}{2\ln 2}} \cdot GB \cdot \left(1 + 2\sqrt{\frac{2\ln 2}{3\pi}} \cdot \frac{F}{B} \cdot \frac{\Delta x}{FWHM}\right) \cdot \frac{FWHM}{\Delta x}$$

$$\sigma^2_{x_c} = [\mathcal{I}_{(x_c,\tilde{F})}(n)^{-1}]_{1,1} \approx \frac{1}{4\ln 2}\sqrt{\frac{\pi}{2\ln 2}} \cdot \frac{B}{GF^2} \cdot \left(1 + \frac{4}{3}\sqrt{\frac{2\ln 2}{3\pi}} \cdot \frac{F}{B} \cdot \frac{\Delta x}{FWHM}\right) \cdot \frac{FWHM^3}{\Delta x}$$

IDS: The Information and Decision Systems Group @ U. de Chile

www.ids.uchile.cl – PI Dr. Jorge Silva

- IDS is a interdisciplinary group focused on applications of estimation and information theory to decision problems in the presence of statistical errors and uncertainty,
- Its stated mission is solving cutting edge inverse problems in the presence of uncertainty by formulating them as instances of estimation or detection tasks,
- IDS is interested in applied research as well as the development of new results and theories.
- Our team has expertise on estimation and decision theory, filtering and prediction, signal processing, information theory and statistical learning,
- The group so far includes Astronomers, Electrical and Electronic Engineers, Geophysicists and Mining Engineers,
  + under- & graduate-students and postdocs.

The Cramér-Rao bound (theorem) in a nutshell

Parameter estimation and the multivariate Cramér-Rao bound

Let I_i (with $i = 1, ..., n$) be a collection of independent observations (or measurements) that follow a parametric probability mass function $f_{\vec{\theta}}$ defined on $\mathbb{N}$. The parameters to be estimated from the measurements $\vec{I} = \{I_i : i = 1, ..., n\}$ will be denoted by $\vec{\theta} = (\theta_1, \theta_2, ..., \theta_m) \in \mathbb{R}^m$. Then given the measurements, let us consider $\hat{\vec{\theta}}(I_1, ..., I_n) = (\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_m)$ to be an unbiased estimator of the parameters $\vec{\theta}$. If $L(\vec{I}; \vec{\theta})$ is the likelihood of the observations $\vec{I}$ given the parameters $\vec{\theta}$, and we can verify that $L(\vec{I}; \vec{\theta})$ satisfies the condition:

$$\mathbb{E}_{\vec{I} \sim f^n_{\vec{\theta}}}\left(\frac{\partial \ln L(\vec{I}; \vec{\theta})}{\partial \theta_i}\right) = 0 \quad \forall \theta_i \;\; (i - 1...m)$$

then, the celebrated Cramér-Rao bound states that (Rao 1945; Cramér 1946):

$$Var(\hat{\theta}_i(\vec{I})) \geq \sigma^2_{\theta_i} \equiv [\mathcal{I}_{\vec{\theta}}(n)^{-1}]_{i,i}$$

where $\mathcal{I}_{\vec{\theta}}(n)$ denotes the *Fisher information* matrix of the data about the vector of parameters $\vec{\theta}$, given by:

$$[\mathcal{I}_{\vec{\theta}}(n)]_{i,j} = \mathbb{E}\left(\frac{\partial \ln L(\vec{I}; \vec{\theta})}{\partial \theta_i} \cdot \frac{\partial \ln L(\vec{I}; \vec{\theta})}{\partial \theta_j}\right) \quad (i, j = 1...m).$$

Application of the CR bound to the Photometry + Astrometry problem on CCD detectors

(Mendez *et al.* 2013, 2014)

$$L(\vec{I}; (x_c, \tilde{F})) = f_{\lambda_1(x_c,\tilde{F})}(I_1) \cdot f_{\lambda_2(x_c,\tilde{F})}(I_2) \cdots f_{\lambda_n(x_c,\tilde{F})}(I_n)$$

$$\ln L(\vec{I}; (x_c, \tilde{F})) = \sum_{i=1}^{n}\left(I_i \cdot \ln \lambda_i(x_c, \tilde{F}) - \lambda_i(x_c, \tilde{F}) - \ln I_i!\right)$$

$$\Downarrow$$

$$\mathcal{I}_{1,1} = \frac{1}{2\pi\sigma^2} \cdot \frac{GF^2}{B} \cdot \sum_{i=1}^{n} \frac{\left(e^{-\gamma(x_i^- - x_c)} - e^{-\gamma(x_i^+ - x_c)}\right)^2}{\left(1 + \frac{1}{\sqrt{2\pi}\,\sigma} \cdot \frac{F}{B} \cdot J_i(x_c)\right)}$$

$$\mathcal{I}_{1,2} = \mathcal{I}_{2,1} = \frac{1}{2\pi\sigma^2} \cdot \frac{F}{B} \cdot \sum_{i=1}^{n} \frac{\left(e^{-\gamma(x_i^- - x_c)} - e^{-\gamma(x_i^+ - x_c)}\right) \cdot J_i(x_c)}{\left(1 + \frac{1}{\sqrt{2\pi}\,\sigma} \cdot \frac{F}{B} \cdot J_i(x_c)\right)}$$

$$\mathcal{I}_{2,2} = \frac{1}{2\pi\sigma^2} \cdot \frac{1}{GB} \cdot \sum_{i=1}^{n} \frac{J_i(x_c)^2}{\left(1 + \frac{1}{\sqrt{2\pi}\,\sigma} \cdot \frac{F}{B} \cdot J_i(x_c)\right)}$$

$f_\lambda(I) = \frac{e^{-\lambda} \cdot \lambda^I}{I!}$

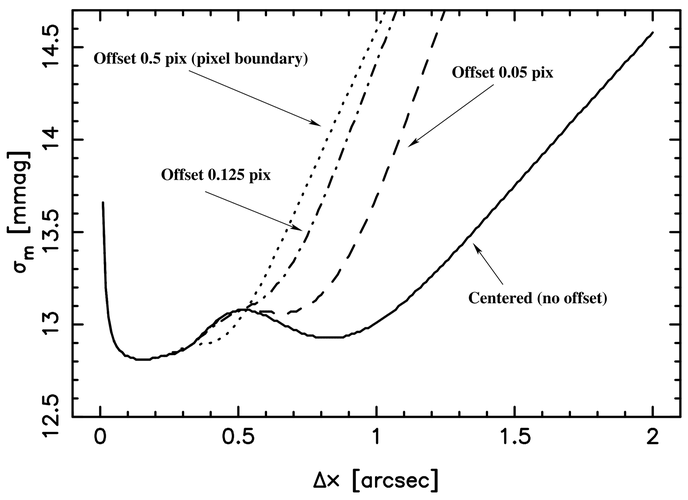
Poisson process

$\lambda_i(x_c, \tilde{F}) = \tilde{F}_i(x_c, \tilde{F}) + \tilde{B}_i$

Expected flux, *per pixel*

$J_i(x_c) \equiv \int_{x_i^-}^{x_i^+} e^{-\gamma(x - x_c)}\, dx \;\; [arcsec]$

Fractional intra-pixel flux (PSF)



Top figure: Photometric CR bound as given by the Fisher matrix coefficients in milli-magnitudes, as a function of detector pixel size Δx in arcsec. The curves were computed for a detector with RON = 5e-, no readout noise, a gain of G = 2e-/ADU, a background of 2,000 ADU/arcsec (representative of ground-based observations on moonless nights through optical broad-band filters with 600 sec exposure time on a good site), and for a Gaussian source with FWHM = 0.5 arcsec and *F* = 5,000 ADU (corresponding to a S/N ~ 80). The solid, dashed, dot-dashed, and dotted lines are for sources which are centered and off-center by 0.05, 0.125 and 0.5 pix (equal to the pixel boundary) respectively. Compare with Figure on the right.

Bottom figure: Astrometric CR bound in milli-arcsec (mas), as a function of detector pixel size Δx in arcsec. All curves were computed for a constant background per pixel *B* = 300 ADU/pix and *F/B* = 10. The dotted, solid, and dashed lines are for a Gaussian PSF with a FWHM of 0.5, 1.0, and 1.5 arcsec respectively, all centered in a given pixel of width 1.0 pixel. The dashed-triple dotted and dashed-dotted lines are for a FWHM of 0.5 arcsec, but de-centered by 0.125 pix and 0.25 pix from the center of the pixel respectively.