CRAMER-RAO BOUNDS AND INSTRUMENT OPTIMIZATION FOR SLITLESS SPECTROSCOPY

Figen S. Oktem[⋆] *Farzad Kamalabadi*[⋆] *Joseph M. Davila*[†]

[⋆] Department of Electrical and Computer Engineering and Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801

[†] NASA Goddard Space Flight Center, Heliophysics Division, Code 671, Greenbelt, MD 20771

ABSTRACT

Spectroscopy is a fundamental diagnostic technique in physical sciences with widespread application. Multi-order slitless imaging spectroscopy has been recently proposed to overcome the limitations of traditional spectrographs, in particular their small instantaneous field of view. Since an inversion is required to infer the physical parameters of interest from slitless spectroscopic measurements, a rigorous theory is essential for quantitative characterization of their performance. In this paper we develop such a theory using the Cramer-Rao lower bounds for the physical parameters of interest, which are derived in terms of important instrument design considerations including the spectral orders to measure, dispersion scale, signal-to-noise ratio, and number of pixels. Our treatment provides a framework for exploring the optimal choices of these design considerations. We illustrate these concepts for an application in EUV solar spectroscopy.

Index Terms— spectroscopy, Cramer-Rao bounds, instrument design, multiframe deblurring

1. INTRODUCTION

Spectroscopy is a fundamental diagnostic technique in physical sciences with application in diverse fields such as physics, chemistry, astronomy, biology, and medicine. Spectrographs enable sensing properties of a scene based on measurement of radiated energy interacting with matter. The measured spectra is the main source of information about the composition and physical properties of targeted objects. For example, in astrophysical imaging of space plasmas, spectral measurements provide estimates of emission line parameters, which are essential for inferring the plasma parameters (such as density, temperature, and flow speed) and understanding the complex plasma behavior [1].

Spectroscopic imaging forms images of a scene at multiple wavelengths. However, obtaining the spectra of a multi-dimensional region with inherently two-dimensional detectors poses intrinsic limitations on the spatio-temporal extent of the technique. Traditionally spectroscopic instruments have suffered from a small instanteneous field-of-view (FOV), i.e. the inability to cover large spatial regions of a scene quickly. A common observation strategy, based on conventional *slit spectrographs*, is to use a narrow field-of-view and a time-consuming rastering process, which as a result limits the spectroscopic analysis of dynamic phenomena. An alternative strategy is to use a *slot spectrograph* by widening the slit; however the blur within the wide slit image, resulting from the dispersion, has then been the bottleneck.

Multi-order slitless spectroscopy has recently been proposed to address these limitations of traditional spectrographs [2]. The problem of estimating spectral line parameters from these spectroscopic measurements can be viewed as a multiframe deblurring problem with shift variant blur, where multiple blurred images of the same object are obtained through multiple spectrometer measurements, each with a different spectral order [3]. However, to date there is no theory that explores the design requirements to reliably infer the physical parameters of interest from these noisy measurements. In this paper we develop such a theory, and, in particular, seek answers to the following questions:

- What minimum data is required for the estimation of physical parameters of interest at a desired precision?
- What is the maximum expected precision in the estimated parameters, and which instrument design considerations can achieve it?
- How much improvement in the precision of estimates is expected with additional spectral orders?

Our tool is the Cramer-Rao lower bound theory [4], which gives the limitations on the precision of unbiased estimates, independent of the estimation method used. We obtain the Cramer-Rao bounds for the estimates of the physical quantities of interest in slitless spectroscopy (for the case of signal dependent Gaussian noise). The Cramer-Rao framework not only allows us to explore the requirements that render this

This work was supported by NASA Headquarters under the NASA Earth and Space Science Fellowship Program - Grant "NNX12AL74H".

new imaging modality effective, but also yields to optimal design choices to minimize the unavoidable precision errors due to noise.

Our approach is similar to some previous works in the literature, which have explored the error bounds for a simpler problem of fitting a single Gaussian line to measurements [5, 6]. The problem in this paper is equivalent to fitting a superposition of multiple Gaussians to the measurements. Recently we have derived approximate analytical bounds for the estimation of spectral line parameters [3]. In this paper, we extend this work to the physical parameters of interest and bring instrument design considerations into the picture.

2. PROBLEM FORMULATION

Recently a novel parametric model for the measurements of multi-order slitless spectroscopy has been developed [7]. Based on this model, each spectroscopic measurement is a blurred version of the same object with a different spatially-varying (Gaussian) filter of some unknown parameters. Let F and Φ^a respectively denote the original image and its blurred version with order a. Considering a row of pixels of length M, the observed intensity of the blurred image in each pixel, Φ^a_m, is related to the original image F, and the blur H^a by

$$\Phi^a_m = \sum_{m'=1}^{M} F_{m'} H^a_{m,m'}. \tag{1}$$

The blur function H^a with $a \neq 0$ is modeled to have a unit Gaussian profile centered at each pixel m', and characterized by a line width, $\Delta_{m'}$, and a line center shift, $\epsilon_{m'}$, which is scaled by the order parameter a:

$$H^a_{m,m'} = \begin{cases} \frac{1}{\Delta_{m'}\sqrt{2\pi}} \exp\left[-\frac{(m-m'-a\,\epsilon_{m'})^2}{2\Delta^2_{m'}}\right] & \text{if } a \neq 0; \\ \delta_{m,m'} & \text{if } a = 0. \end{cases} \tag{2}$$

Note that when the order is zero, there is no blur on the image, i.e. $\Phi^0_m = F_m$. Here Φ^a, F, ϵ, and Δ all have the size M, and both line widths Δ and center shifts ϵ are measured in pixel units.

We consider the following noise model for the measurements:

$$\tilde{\Phi}^a_m = \Phi^a_m + V^a_m, \quad V^a_m \sim N(0, \left(\frac{\Phi^a_m}{SNR}\right)^2) \tag{3}$$

where we assume that the observation noise V^a_m is signal-dependent Gaussian noise that is uncorrelated across both m and a. This noise model is also a good approximation to signal-dependent Poisson noise when the signal values are sufficiently large. Here the noise variance is expressed in terms of the signal-to-noise ratio, SNR, which is defined as the ratio of the signal mean to the standard deviation of the noise.

In the intended application of space plasma imaging, an observed blurred image $\tilde{\Phi}^a$ corresponds to a dispersed image with spectral order a and the parameters F, Δ, and ϵ respectively model the integrated intensities, widths, and Doppler shifts of the emission lines radiated from each pixel.

3. CRAMER-RAO ERROR BOUNDS

In the inverse problem, the goal is to estimate the unknown parameters F, Δ, and ϵ (i.e. the spectral line parameters) from the multiple measurements of Φ^a with orders $a \in \mathcal{A}$. (This problem can be viewed as a multi-frame deblurring problem with shift-variant blur [3].) By Cramer-Rao lower bound (CRB) theory [4], the variance of unbiased estimates of the parameters $\{F, \Delta, \epsilon\}$ provided by any method is always lower bounded by

$$var(\hat{\theta}_i) \geq [I(\theta)^{-1}]_{ii}. \tag{4}$$

where θ denotes the parameters $\{F, \Delta, \epsilon\}$, $\hat{\theta}_i$ denotes any unbiased estimator of the parameter θ_i (corresponding to the ith element of θ), and $I(\theta)$ is the Fisher information matrix whose (i,j)th element is given by

$$[I(\theta)]_{ij} = -E\left[\frac{\partial^2}{\partial\theta_i\,\partial\theta_j}\log L\right] \tag{5}$$

where the right side is evaluated at the true value of the parameters, denoted by θ. In our problem, the log-likelihood function, $\log L$, is [3]

$$\log L \propto -\frac{1}{2}\sum_{m=1}^{M}\sum_{a\in\mathcal{A}}\left(\frac{SNR}{\Phi^a_m}\right)^2(\tilde{\Phi}^a_m - \Phi^a_m)^2 - \log\left(\frac{\Phi^a_m}{SNR}\right)$$

where the constant term that is independent of the unknown parameters is omitted. Hence the elements of the Fisher information matrix can be obtained as

$$\begin{aligned} [I(\theta)]_{ij} &= (SNR^2+2)\sum_{m=1}^{M}\sum_{a\in\mathcal{A}}\frac{1}{(\Phi^a_m)^2}\frac{\partial\Phi^a_m}{\partial\theta_i}\frac{\partial\Phi^a_m}{\partial\theta_j} \\ &= (SNR^2+2)\left(\sum_{a\in\mathcal{A}} J_a^T \Sigma^2_{\Phi^a} J_a\right)_{ij} \end{aligned} \tag{6}$$

where $\Sigma_{\Phi^a} = diag(1/\Phi^a_1 \ldots 1/\Phi^a_M)$ is a diagonal matrix, and J_a is $M \times 3M$ Jacobian matrix of Φ_a, that is $(J_a)_{mi} = \frac{\partial\Phi^a_m}{\partial\theta_i}$. The Jacobian matrices J_a with $a \neq 0$ can be computed from the first derivatives of the blurred images as follows:

$$\begin{aligned} (J_a)_{m,3k+1} &= \frac{1}{\Delta_k\sqrt{2\pi}} e^{-\frac{(m-k-a\,\epsilon_k)^2}{2\Delta^2_k}} \\ (J_a)_{m,3k+2} &= \frac{F_k}{\Delta^2_k\sqrt{2\pi}} e^{-\frac{(m-k-a\,\epsilon_k)^2}{2\Delta^2_k}} \left(\frac{(m-k-a\,\epsilon_k)^2}{\Delta^2_k} - 1\right) \\ (J_a)_{m,3k+3} &= \frac{F_k\,a}{\Delta^3_k\sqrt{2\pi}} e^{-\frac{(m-k-a\,\epsilon_k)^2}{2\Delta^2_k}} (m-k-a\,\epsilon_k) \end{aligned}$$

whereas for $a = 0$, $(J_0)_{m,3k+1} = \delta_{m,k}$, and $(J_0)_{m,3k+2} = (J_0)_{m,3k+3} = 0$, for all m and k. Then the Cramer-Rao error bounds for each parameter can be obtained from the diagonal elements of the inverse Fisher information matrix.

4. ERROR BOUNDS FOR PHYSICAL QUANTITIES

The error bounds derived in the previous section are for the spectral line parameters F_k, Δ_k, and ϵ_k, where both Δ_k and ϵ_k are measured in pixels. However the goal in spectroscopic imaging is to estimate some physical quantities of interest related to these spectral line parameters. Hence one would generally be interested in the error bounds for these physical parameters (measured in physical units), rather than for the spectral line parameters (measured in pixels). Moreover the goal in instrument optimization is generally to minimize the errors for some *physical parameters* with respect to *design parameters*. These require us to express the spectral line parameters in terms of the physical quantities of interest and design parameters.

We start by relating Δ_k and ϵ_k to the physical quantities they represent, using the dispersion scale D. Dispersion scale (also known as reciprocal dispersion) is the wavelength range corresponding to a single pixel, and is measured here in mÅ/pixel. (Low dispersion scale means large dispersion in the instrument.) The primary physical quantities of interest w_k and v_k, respectively denoting the line width in wavelength units and the line-of-sight velocity, can then be expressed as

$$w_k = D\,\Delta_k \;\Rightarrow\; \Delta_k = \frac{w_k}{D} \tag{7}$$

and

$$v_k = \frac{c(D\epsilon_k)}{1000\lambda_0} \;\Rightarrow\; \epsilon_k = \frac{1000\lambda_0 v_k}{cD} \tag{8}$$

Here w_k is measured in mÅ, v_k is measured in km/s, c denotes the speed of light in km/s, and λ_0 represents the central wavelength in Å. (The second relation is obtained using the Doppler shift formula: $\frac{\Delta\lambda}{\lambda_0} = \frac{v}{c}$. This gives the relation between the line-of-sight velocity v and the resulting wavelength shift $\Delta\lambda$ of the central wavelength λ_0.)

The error bounds for F_k, w_k, v_k can be obtained from the error bounds for F_k, Δ_k, ϵ_k by using the above relations and performing a transformation of the parameters. Let the new parameter set be $\theta' = (F_k, w_k, v_k)$. With the old parameter set being $\theta = (F_k, \Delta_k, \epsilon_k)$, the relation between these two parameter sets is given by

$$\theta = (F_k, \frac{w_k}{D}, \frac{1000\lambda_0}{Dc} v_k) = g(\theta'). \tag{9}$$

Then the Fisher information matrix for the new parameter set $\theta' = (F_k, w_k, v_k)$ can be obtained by using the error propagation formula [8] as follows:

$$I(\theta') = \left(\frac{\partial g}{\partial \theta'}\right)^T I(g(\theta')) \frac{\partial g}{\partial \theta'} \tag{10}$$

where $I(g(\theta'))$ is the Fisher information matrix for the old parameter set (given in (6)), and the Jacobian matrix $\frac{\partial g}{\partial \theta'}$ is a block diagonal matrix with each block given by

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{D} & 0 \\ 0 & 0 & \frac{1000\lambda_0}{Dc} \end{bmatrix}.$$

A few remarks are in order at this point. First, the new Fisher information matrix and the corresponding error bounds for the physical parameters are functions of the instrument design choices, consisting of the spectral orders to be measured, SNR level, dispersion scale, and the number of pixels. All of these design choices are significant in determining the amount of information available from the slitless data, and their effect on the precision of estimates can be explored using the error bounds. Second, the given formulation is general enough to allow consideration of other physical quantities of interest and design parameters.

5. NUMERICAL EXAMPLE

Thus far, the analytic framework developed is quite general and can be applied to any particular application involving spectroscopy. In this section, we illustrate the usefulness of these concepts for an application in EUV solar spectroscopy.

We consider the measurements of an EUV emission line emitted from the Sun, with a central wavelength of $\lambda_0 = 335$ Å. Our goal is to explore the effectiveness of slitless spectroscopy to this application, and to find the optimal design choices for the spectral orders, dispersion scale, SNR, and number of pixels. The optimality criteria should be chosen based on the science objectives of the experiments performed with the instrument. As an example, we consider here a commonly used criterion: the one that minimizes the average of standard deviations for F_k, w_k, and v_k. We compute this optimality criterion as a function of one design parameter while fixing the others.

Various observing scenarios with three or more spectral orders are considered in the simulations. The cases with less than three orders are not included because in these cases the Fisher information matrix is observed to be highly ill-conditioned. This indicates that these measurement configurations suffer from instability issues [9], and hence cannot yield stable estimates of the physical parameters. In our analysis we only consider the instances with well-conditioned Fisher information matrix so that we can reliably invert the matrix to compute the error bounds.

The spectral line parameters are generated randomly according to their modeled distributions in solar spectroscopy. Although the focus of this paper is CRB for unbiased estimators, our next intent is to extend these bounds for maximum a posteriori estimators. For this reason, the numerical averages of the bounds are computed for a total of 50 random parameter sets. Figures 1, 2 and 3 show these results.

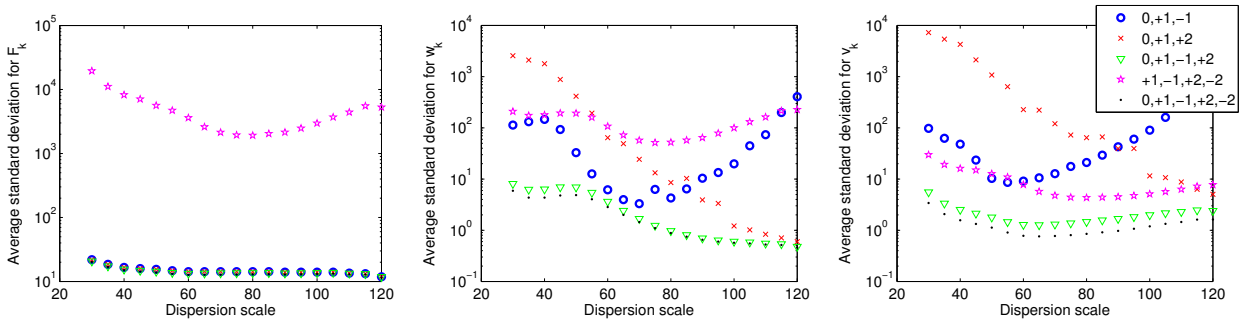

Fig. 1: Average standard deviations for F_k, w_k, and v_k as a function of the dispersion scale with $SNR = 100$ and $M = 100$.

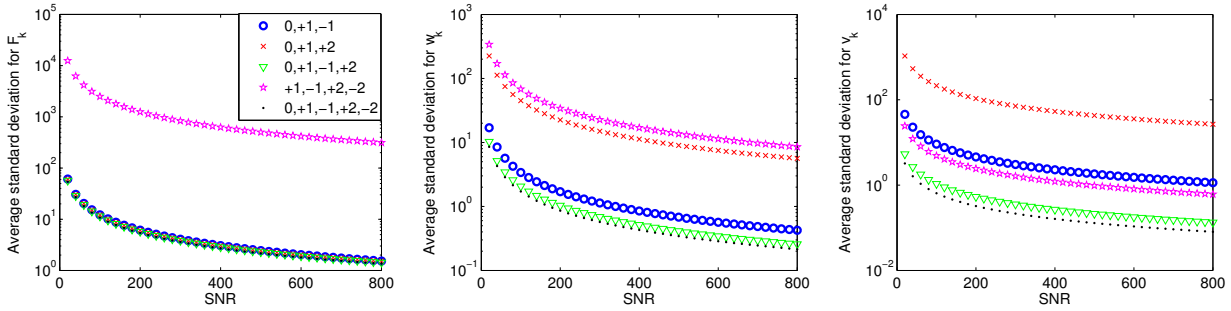

Fig. 2: Average standard deviations for F_k, w_k, and v_k as a function of SNR with $D = 65$ mÅ/pix and $M = 100$.

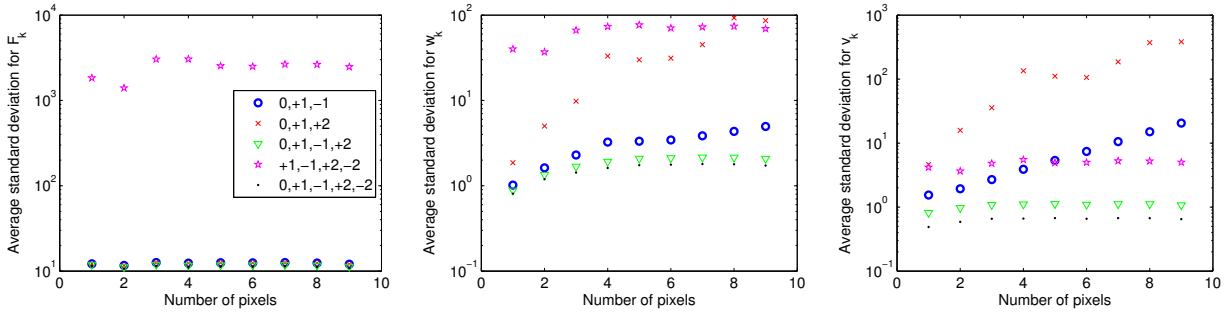

Fig. 3: Average standard deviations for F_k, w_k, and v_k as a function of number of pixels with $D = 65$ mÅ/pix and $SNR = 100$.

The plots provide important insights about how to optimally operate this imaging spectroscopy. Here we note the most important observations. First of all, optimal dispersion range is similar for both the line widths and line-of-sight velocities, and the weak trade-off can be handled depending on the science goal. For integrated intensities, the error bounds depend only on whether the 0th spectral order is measured or not, and the SNR of this measurement. Furthermore, from the expression of the Fisher information matrix, we already know that the rms errors are proportional to $\sim 1/SNR$. From the plots, we observe that SNR values of larger than 200 yield only slight improvements in errors.

When we compare the various observing scenarios with different orders, the case with orders $\{0, +1, -1\}$ operated at a nearly-optimal dispersion appears to be the most cost-effective one. In this case, additional fourth and fifth orders provide almost no improvement for estimating integrated intensities and line widths; but the additional fourth order yields some improvement in velocity estimates. If operated in an optimal regime, the performance of the slitless spectrograph is comparable to the conventional method of slit spectrograph (in terms of precision of the estimates), but with the additional advantage of a large instantenous FOV.

6. REFERENCES

[1] K. J. H. Phillips, U. Feldman, and E. Landi, *Ultraviolet and X-ray Spectroscopy of the Solar Atmosphere*, Cambridge Univ. Press, 2008.

[2] C. C. Kankelborg and R. J. Thomas, "Simultaneous imaging and spectroscopy of the solar atmosphere: advantages and challenges of a 3-order slitless spectrograph," in *Proc. SPIE UV/EUV and Visible Space Instrumentation for Astronomy and Solar Physics*, vol. 4498, pp. 16–26, 2001.

[3] F. S. Oktem and F. Kamalabadi, "Analytical precision limits in slitless spectroscopy," in *Proc. 2012 IEEE Statistical Signal Processing Workshop*, pp. 468–471, 2012.

[4] S. M. Kay, *Fundamentals of Statistical Signal Processing*, Prentice Hall, 1993.

[5] J. Ireland, "Precision limits to emission-line profile measuring experiments," *Astrophys. J.*, vol. 620, pp. 1132-1139, 2005.

[6] S. Minin and F. Kamalabadi, "Uncertainties in extracted parameters of a gaussian emission line profile with continuum background," *Applied optics*, vol. 48, no. 36, pp. 6913–6922, 2009.

[7] J. M. Davila, "Slitless solar spectroscopy," in *AAS/Solar Physics Division Meeting*, vol. 42, 2011.

[8] E.L. Lehmann and G. Casella, *Theory of point estimation*, vol. 31, Springer, 1998.

[9] S. Basu and Y. Bresler, "The stability of nonlinear least squares problems and the cramer-rao bound," *IEEE Trans. Signal Processing*, vol. 48, no. 12, pp. 3426–3436, 2000.