

# Bias-Resolution-Variance Tradeoffs for Single Pixel Estimation Tasks Using the Uniform Cramér-Rao Bound<sup>1</sup>

Thomas J. Kragh<sup>1</sup> and Alfred O. Hero, III<sup>1</sup>

<sup>1</sup>University of Michigan, Department of EECS, Ann Arbor, MI, 48109, USA  
tkragh@eeecs.umich.edu, hero@eeecs.umich.edu

## Abstract

Previously we introduced the Uniform Cramér-Rao (CR) Bound as a lower bound on the variance of biased estimators, along with the concept of the delta-sigma tradeoff curve. For an estimator whose variance lies on this curve, lower variance can only be achieved at the price of increased estimator bias gradient norm, and vice versa.

However, for single pixel estimation, one can specify a variety of different estimator point response functions that have identical bias-gradient norm but with widely different resolution properties. This has led to some counter-intuitive results and interpretation difficulties when using the Uniform CR Bound in performance studies of imaging systems.

We now extend this tradeoff concept by introducing the 2nd-moment of the point response function as a measure of resolution for single-pixel estimation tasks. We derive an expression for the delta-gamma-sigma tradeoff surface. This surface specifies an "unachievable region" of estimator variance. For estimators that lie on this surface, lower variance can only be achieved at the price of increased bias gradient norm and/or decreased estimator resolution. We present a method for computing this surface for linear gaussian inverse problems.

## I. INTRODUCTION

We previously introduced a method for specifying a lower bound on the variance of biased estimators using the Uniform Cramér-Rao (CR) Bound, along with the concept of the delta-sigma tradeoff curve [1]. For an estimator whose variance lies on this curve, lower estimator variance can only be achieved at the price of an increased estimator bias gradient norm, and vice versa.

The Uniform CR Bound has been used to calculate fundamental limits in estimator performance in medical imaging [2-3], comparing the performance of different medical imaging modalities [4], among other applications.

In [1], we showed that a) the estimator bias gradient norm is an upper bound on the maximal squared variation of the estimator bias function over an ellipsoidal neighborhood, and b) equivalent to the difference between the estimator mean response to a point source, and the true response. Thus, the norm of the estimator bias gradient would seem to be a natural measure of estimator bias.

One problem with using bias gradient norm when comparing the variance of different estimators and/or systems is that different estimator point response functions can have identical bias-gradient norm but widely different resolution

properties. This has led to interpretation difficulties and counter-intuitive results when using the Uniform CR Bound in imaging system performance studies. Figure 1 shows three example point response functions with similar FWHM and identical bias gradient length, yet with obviously different resolution properties.

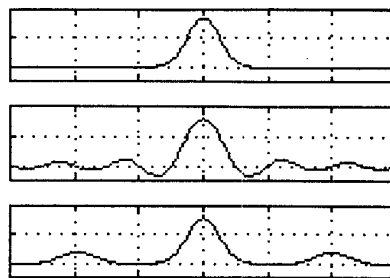


Figure 1: Three example point response functions with identical bias gradient norms and similar FWHM.

In this paper, we will introduce a fundamental tradeoff relationship between bias, resolution, and variance for single pixel estimation. Along with the estimator bias gradient norm, we now introduce the 2<sup>nd</sup>-moment of the estimator point response function as a resolution measure. A Uniform CR Bound will be derived for the variance of single-pixel estimators as a function of both the estimator bias gradient norm and 2nd-moment of the point response function. The concept of tradeoffs in estimator variance now include both overall bias error (as measured by the bias gradient norm) along with resolution error (as measured by the 2nd-moment of the estimator point response function). The surface parameterized by bias gradient norm and 2nd-moment specifies an "unachievable region" of estimator variance. For estimators that lie on this surface, lower variance can only be achieved at the price of increased bias gradient norm and/or decreased estimator resolution as measured by the 2nd-moment of the estimator point response.

## II. DEFINITIONS

### A. Statistical Model

Let  $\underline{\theta} = [\theta_1, \dots, \theta_n]^T \in \Theta$  be a vector of unknown, nonrandom parameters that parameterize the density  $f_Y(y|\underline{\theta})$

of the measured random variable  $\mathbf{Y}$ . The parameter space  $\Theta$  is assumed to be an open subset of the  $n$ -dimensional Euclidean space  $R^n$ . Let  $\hat{\theta}_p = \hat{\theta}_p(\mathbf{Y})$  be an estimator of the  $p^{\text{th}}$ -component of  $\underline{\theta}$ . Let this estimator have mean value  $m_{\theta} = E_{\theta}[\hat{\theta}_p]$ , bias  $b_{\theta} = E_{\theta}[\hat{\theta}_p] - \theta_p$ , and variance  $\sigma_{\theta}^2 = E_{\theta}[(\hat{\theta}_p - \theta_p)^2]$ . The estimator  $\hat{\theta}_p$  can be expressed in terms of the vector parameter estimator  $\hat{\theta} = \hat{\theta}(\mathbf{Y})$  via  $\hat{\theta}_p = \underline{e}_p^T \hat{\theta}$ , where  $\underline{e}_p = (0, \dots, 0, 1, 0, \dots, 0)^T$  (the  $p^{\text{th}}$ -unit vector). The gradient of the estimator bias function  $\nabla b_{\theta}$  and mean response gradient  $\nabla m_{\theta}$  are therefore related by  $\nabla m_{\theta} = \underline{e}_p + \nabla b_{\theta}$ .

In [1], we showed that under certain conditions for estimating the  $p^{\text{th}}$ -pixel in both the linear Gaussian and nonlinear Poisson inverse problems, the bias gradient vector  $\nabla b_{\theta}$  is equivalent to the difference between the mean of the vector estimator response  $E_{\theta}[\hat{\theta}]$  to a point source  $\underline{e}_p$ , and the true point response  $\underline{e}_p$ . Similarly, the mean of the gradient of the  $p^{\text{th}}$ -pixel estimator  $\nabla m_{\theta}$  is equivalent to the mean estimator  $E_{\theta}[\hat{\theta}]$  when the input is a point source.

### B. Overall Bias and Resolution Measures

We will define the bias gradient norm  $\delta$  with respect to a positive definite matrix  $C$ , and the point response  $2^{\text{nd}}$ -moment  $\gamma$  as

$$\delta^2 = \nabla b_{\theta}^T C \nabla b_{\theta} \quad (1)$$

$$\gamma^2 = \frac{\sum_i (p-i)^2 (\nabla m_{\theta})_i^2}{\sum_i (\nabla m_{\theta})_i^2} \quad (2)$$

Since the mean-response gradient is the sum of the true point-response and bias-gradient, Equation #2 can be rewritten as the ratio of two quadratic forms,

$$\gamma^2 = \frac{(\underline{e}_p + \nabla b_{\theta})^T M_p (\underline{e}_p + \nabla b_{\theta})}{(\underline{e}_p + \nabla b_{\theta})^T (\underline{e}_p + \nabla b_{\theta})} \quad (3)$$

where  $M_p$  is a positive semi-definite diagonal matrix with diagonal elements proportional to the square of the distance of each pixel from the  $p^{\text{th}}$ -pixel. For the 1D case,  $M_p$  is

$$[M_p]_{ij} = (p-i)^2 \delta(i-j) \quad (4)$$

where  $\delta(i-j)$  is (for this expression only) the discrete delta function.

### C. CR Bound for Biased Estimators

For a biased estimator  $\hat{\theta}_p$  of the  $p^{\text{th}}$ -pixel, the biased estimator CR bound of estimator variance is given by

$$\sigma_{\theta}^2 \geq [\underline{e}_p + \nabla b_{\theta}]^T F_Y^+ [\underline{e}_p + \nabla b_{\theta}] \quad (5)$$

where the  $n \times n$  Fisher Information matrix  $F_Y$  is given by

$$F_Y = E_{\theta} \left\{ \left[ \nabla_{\theta} \ln(f_Y(\mathbf{y}|\theta)) \right] \left[ \nabla_{\theta} \ln(f_Y(\mathbf{y}|\theta)) \right]^T \right\} \quad (6)$$

and  $F_Y^+$  is the Moore-Penrose pseudo-inverse of the (possibly singular) Fisher Information Matrix.

Although (5) does give a lower bound on the variance of a biased estimator, it only applies to estimators with bias gradient vector  $\nabla b_{\theta}$ . Since the bias-gradient is typically unique to a particular estimator, equation (5) will be estimator specific.

### III. UNIFORM CR-BOUND

In order to have a bound on estimator variance that does not depend on the specific bias gradient  $\nabla b_{\theta}$ , the Uniform CR Bound presented in [1] gave a lower bound to equation (5) subject to a constraint on the *bias-gradient length* or *norm*  $[\nabla b_{\theta}]_C^2 = \nabla b_{\theta}^T C \nabla b_{\theta}$ , where  $C$  is a positive-definite matrix.

We note that the estimator mean response to a point source  $E_{\theta}[\hat{\theta}]$  is equivalent to the gradient of the mean estimator response  $\nabla m_{\theta}$ , and that  $\nabla m_{\theta} = \underline{e}_p + \nabla b_{\theta}$ . We will now add an additional constraint on the estimator bias gradient. Since  $\nabla m_{\theta}$  is equivalent to  $E_{\theta}[\hat{\theta}]$ , and  $\nabla m_{\theta} = \underline{e}_p + \nabla b_{\theta}$ , a constraint on the  $2^{\text{nd}}$ -moment of the estimator response to a point source can be expressed as an additional constraint on  $\nabla b_{\theta}$ .

#### A. UCRB with Bias Gradient and Resolution Constraints

Here we present a Uniform CR Bound for biased estimators with a given bias-gradient length  $\delta$  and  $2^{\text{nd}}$ -moment measure  $\gamma$  for a non-singular Fisher Information matrix  $F_Y$ .

Let  $\hat{\theta}_p$  be an estimator of the  $p^{\text{th}}$ -component of the parameter vector  $\underline{\theta}$ . For a fixed  $\delta, \gamma \geq 0$ , let the bias gradient satisfy the constraints  $\nabla b_{\theta}^T C \nabla b_{\theta} \leq \delta^2$  and

$\frac{(\underline{e}_p + \nabla b_{\theta})^T M_p (\underline{e}_p + \nabla b_{\theta})}{(\underline{e}_p + \nabla b_{\theta})^T (\underline{e}_p + \nabla b_{\theta})} \leq \gamma^2$ . Then the variance  $\sigma_{\theta}^2$  of  $\hat{\theta}_p$  satisfies

$$\sigma_{\theta}^2 \geq B(\underline{e}, \delta, \gamma) \quad (7)$$

where the value of  $B(\underline{\theta}, \delta, \gamma)$  is given by the following three cases:

I) If  $\delta^2 \geq \nabla b_{\theta}^T C \nabla b_{\theta}$ , then

$$B(\underline{\theta}, \delta, \gamma) = 0 \quad (8)$$

II) If  $\delta^2 < \nabla b_{\theta}^T C \nabla b_{\theta}$  and  $\gamma \geq \gamma^*$ , then

$$B(\underline{\theta}, \delta, \gamma) = \lambda_1^2 \underline{\epsilon}_p^T [\lambda_1 F_Y + C^{-1}]^{-1} F_Y [\lambda_1 F_Y + C^{-1}]^{-1} \underline{\epsilon}_p \quad (9)$$

where  $\lambda_1 > 0$  is a lagrange multiplier given by the unique solution to

$$\delta^2 = \underline{\epsilon}_p^T [\lambda_1 F_Y + C^{-1}]^{-1} C^{-1} [\lambda_1 F_Y + C^{-1}]^{-1} \underline{\epsilon}_p \quad (10)$$

and

$$\gamma^* = \frac{(\underline{\epsilon}_p + \underline{d}_{\min}^*)^T M_p (\underline{\epsilon}_p + \underline{d}_{\min}^*)}{(\underline{\epsilon}_p + \underline{d}_{\min}^*)^T (\underline{\epsilon}_p + \underline{d}_{\min}^*)} \quad (11)$$

and

$$\underline{d}_{\min}^* = -C^{-1} [C^{-1} + \lambda_1 F_Y]^{-1} \underline{\epsilon}_p \quad (12)$$

III) If  $\delta^2 < \nabla b_{\theta}^T C \nabla b_{\theta}$  and  $\gamma < \gamma^*$ , then

$$B(\underline{\theta}, \delta, \gamma) = [\underline{\epsilon}_p + \underline{d}_{\min}] F_Y^{-1} [\underline{\epsilon}_p + \underline{d}_{\min}] \quad (13)$$

where

$$\underline{d}_{\min} = -[F_Y^{-1} + \lambda_1 C + \lambda_2 [M_p - \gamma^2 I]]^{-1} [F_Y^{-1} - \lambda_2 \gamma^2 I] \underline{\epsilon}_p \quad (15)$$

and  $\lambda_1, \lambda_2 \geq 0$  are lagrange multipliers found implicitly through the two equality constraints

$$(\underline{d}_{\min} + \underline{\epsilon}_p)^T [M_p - \gamma^2 I] (\underline{d}_{\min} + \underline{\epsilon}_p) = 0 \quad (16)$$

$$\underline{d}_{\min}^T C \underline{d}_{\min} - \gamma^2 = 0 \quad (17)$$

## IV. RESULTS

### A. Linear Gaussian Model

We generated a delta-gamma-sigma surface to show the lower bound of a single pixel estimator for a 1D deconvolution problem. The measurement equation is  $\underline{Y} = A\underline{\theta} + \underline{\epsilon}$ . The parameter and measurement vectors are both 128 element vectors corresponding to a 128 pixel 1D signal. The system matrix A corresponds to a shift-invariant gaussian shaped blur kernel with a full-width-half-max of 1.0 pixels. The additive noise vector  $\underline{\epsilon}$  is distributed normally with mean 0 and covariance  $\Sigma = I$ . The Fisher Information Matrix is  $F_Y = A^T \Sigma^{-1} A$ . The estimation task is for the 67<sup>th</sup> pixel, thus  $\hat{\theta}_{\rho}(\underline{Y}) = \hat{\theta}_{67}(\underline{Y})$ .

We also overlay on the surface the variance of two different Penalized Weighted Least Squares (PWLS)

estimators  $\hat{\theta}(\underline{Y}) = [F_Y + \beta P]^{-1} A^T \Sigma^{-1} \underline{Y}$ . The first estimator has penalty matrix  $P$  equal to the 1<sup>st</sup>-order neighborhood difference (Laplacian) matrix, and the second estimator has penalty  $P$  equal to the Identity matrix. Note that the first estimator penalty corresponds to a roughness penalty, whereas the second estimator penalty corresponds to an energy penalty or Tikonov regularization. Both estimators have a regularization parameter  $\beta \geq 0$ . The measurement covariance used in the estimators was purposely set at 2x the true covariance, to show the estimator trajectories lying above the bound surface. The choice of norm matrix is  $C = P^{-1}$ .

Figure 2 shows a plot of this surface, along with the trajectories in (bias, resolution, variance) of the two PWLS estimators for values of the penalty  $10^{-3} \leq \beta \leq 10^4$ .

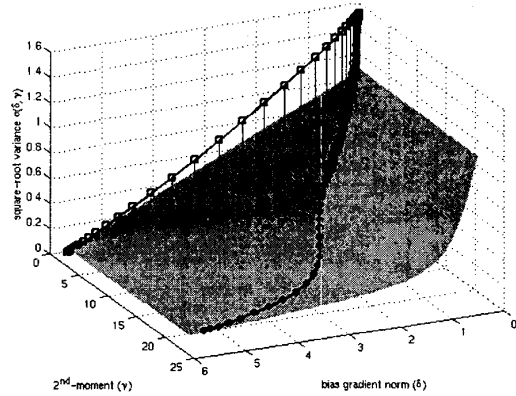


Figure 2: delta-gamma-sigma surface for linear Gaussian inverse problem, along with trajectories of PWLS estimators with roughness penalty (circular ticks) and identity penalty (square ticks).

## V. REFERENCES

- [1] A.O. Hero, J.A. Fessler, and M. Usman, "Exploring Estimator Bias-Variance Tradeoffs Using the Uniform CR Bound" IEEE Trans. Signal Processing, vol. 44, pp.2026-2041, August 1996.
- [2] M. Usman, A.O. Hero and J. A. Fessler, "Uniform CR bound: implementation issues and applications to image reconstruction", Proceedings of the IEEE 1994 NSS-MIC, Norfolk VA, Oct. 1994.
- [3] Chor-yi Ng, N. Clinthorne, J.A. Fessler, M. Usman, A.O. Hero, and W.L. Rogers, "Preliminary studies on the feasibility of addition of a vertex view to conventional brain SPECT", Proceedings of the 1996 IEEE NSS-MIC, Anaheim CA, Nov. 1996, pp. 1561-1565
- [4] N.H. Clinthorne, C-y Ng, C-h. Hua, J.E. Gormley, J.W. LeBlanc, D.K. Wehe, S.J. Wilderman, and W.L. Rogers, "Theoretical performance comparison of a Compton scatter aperture and parallel hole collimator", Proceedings of the 1996 IEEE NSS-MIC, Anaheim CA, Nov 1996 pp. 788-792