Bias-Resolution-Variance Tradeoffs for Single Pixel Estimation Tasks using the Uniform Cramér-Rao Bound

Thomas J. Kragh, and Alfred O. Hero, III

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 estimators whose variance lie 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 lead to some counter-intuitive results and interpretation difficulties when using the Uniform CR Bound in performance studies of imaging systems. In this paper, we 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 and nonlinear Poisson inverse problems. Finally, we will show bound calculations for a Compton-SPECT imaging system.

T. J. Kragh and A. O. Hero are with the Department of Electrical Engineering, The University of Michigan, Ann Arbor, MI 48109 (email: tkragh@eecs.umich.edu, hero@eecs.umich.edu)

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 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.

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 norms but widely different resolution properties. This has lead 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.

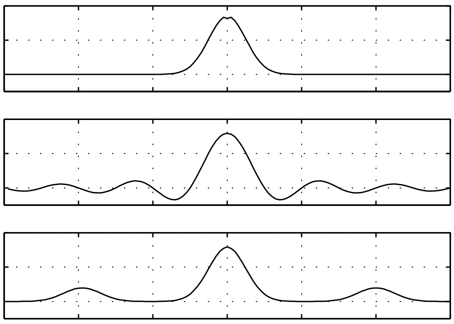

Figure 1: Three example point response functions with identical bias gradient norms and similar FWHM.

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 2nd-moment of the estimator point response function as a resolution measure. A Uniform CR Bound will then 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). This 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 $\underline{\theta} = [\theta_1, ..., \theta_n]^T \in \Theta$ be a vector of unknown, nonrandom parameters that parameterize the density $f_Y(y|\underline{\theta})$ of the measured random variable Y. The parameter space Θ is assumed to be an open subset of the n-dimensional Euclidean space R^n. Let $\hat{\theta}_p = \hat{\theta}_p(Y)$ be an estimator of the p^{th}-component of $\underline{\theta}$. Let this estimator have mean value $m_{\underline{\theta}} = E_{\underline{\theta}}[\hat{\theta}_p]$, bias $b_{\underline{\theta}} = E_{\underline{\theta}}[\hat{\theta}_p] - \theta_p$, and variance $\sigma_{\underline{\theta}}^2 = E_{\underline{\theta}}\left[\left(\hat{\theta}_p - \theta_p\right)^2\right]$. The estimator $\hat{\theta}_p(Y)$ can be expressed in terms of the vector parameter estimator $\underline{\hat{\theta}}(Y)$ via $\hat{\theta}_p = \underline{e}_p^T \underline{\hat{\theta}}$, where $\underline{e}_p = [0, ..., 0, 1, 0, ..., 0]^T$ (the p^{th}-unit vector). The bias gradient and mean response gradient of $\hat{\theta}_p(Y)$ are therefore related by $\nabla m_{\underline{\theta}} = \underline{e}_p + \nabla b_{\underline{\theta}}$,

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

B. Overall Bias and Resolution Measures

We will define the bias gradient norm δ with respect to a positive definite matrix C, along with the point response 2nd-moment γ as

$$\delta^2 = \nabla b_{\underline{\theta}}^T C \nabla b_{\underline{\theta}} \tag{1}$$

$$\begin{aligned}\gamma^2 &= \frac{\sum_i (p-i)^2 \left(\nabla m_{\underline{\theta}}\right)_i^2}{\sum_i \left(\nabla m_{\underline{\theta}}\right)_i^2} \\ &= \frac{\nabla m_{\underline{\theta}}{}^T M_p \nabla m_{\underline{\theta}}}{\nabla m_{\underline{\theta}}{}^T \nabla m_{\underline{\theta}}} \\ &= \frac{\left(\underline{e}_p + \nabla b_{\underline{\theta}}\right)^T M_p \left(\underline{e}_p + \nabla b_{\underline{\theta}}\right)}{\left(\underline{e}_p + \nabla b_{\underline{\theta}}\right)^T \left(\underline{e}_p + \nabla b_{\underline{\theta}}\right)}\end{aligned} \tag{2}$$

For a 1-D imaging system, 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^{th}-pixel

$$\left[M_p\right]_{ij} = (p-i)^2 \delta(i-j) \tag{3}$$

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^{th}-pixel value, the biased CR bound of estimator variance is given by

$$\sigma_{\underline{\theta}}^2 \ge \left[\underline{e}_p + \nabla b_{\underline{\theta}}\right]^T F_Y^+ \left[\underline{e}_p + \nabla b_{\underline{\theta}}\right] \tag{4}$$

where the n x n Fisher Information matrix F_Y is

$$F_Y = E_{\underline{\theta}}\left\{\left[\nabla_{\underline{\theta}} \ln\left(f_Y\left(y|\underline{\theta}\right)\right)\right]\left[\nabla_{\underline{\theta}} \ln\left(f_Y\left(y|\underline{\theta}\right)\right)\right]^T\right\} \tag{5}$$

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

III. Uniform CR Bound

Here we present a Uniform CR Bound for non-singular Fisher Information matrix F_Y, the proof of which will be given in the appendix section. Let $\hat{\theta}_p$ be an estimator of the p[th]-component of the parameter vector $\underline{\theta}$. For a fixed $\delta, \gamma \ge 0$, let the bias gradient satisfy the constraints $\nabla b_{\underline{\theta}}^T C \nabla b_{\underline{\theta}} \le \delta^2$ and $\frac{\left(\underline{e}_p + \nabla b_{\underline{\theta}}\right)^T M_p \left(\underline{e}_p + \nabla b_{\underline{\theta}}\right)}{\left(\underline{e}_p + \nabla b_{\underline{\theta}}\right)^T \left(\underline{e}_p + \nabla b_{\underline{\theta}}\right)} \le \gamma^2$. Then the variance $\sigma_{\underline{\theta}}^2$ of $\hat{\theta}_p$ satisfies

$$\sigma_{\underline{\theta}}^2 \ge B\left(\underline{\theta}, \delta, \gamma\right) \tag{6}$$

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

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

$$B\left(\underline{\theta}, \delta, \gamma\right) = 0 \tag{7}$$

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

$$B\left(\underline{\theta}, \delta, \gamma\right) = \lambda_1^2 \underline{e}_p^T \left[\lambda_1 F_Y + C^{-1}\right]^{-1} F_Y \left[\lambda_1 F_Y + C^{-1}\right]^{-1} \underline{e}_p \tag{8}$$

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

$$\delta^2 = g\left(\lambda_1\right) = \underline{e}_p^T \left[\lambda_1 F_Y + C^{-1}\right]^{-1} C^{-1} \left[\lambda_1 F_Y + C^{-1}\right]^{-1} \underline{e}_p \tag{9}$$

and

$$\gamma^* = \frac{\left(\underline{e}_p + \underline{d}_{\min}^*\right)^T M_p \left(\underline{e}_p + \underline{d}_{\min}^*\right)}{\left(\underline{e}_p + \underline{d}_{\min}^*\right)^T \left(\underline{e}_p + \underline{d}_{\min}^*\right)} \tag{10}$$

and $$\underline{d}_{\min}^* = -C^{-1}\left[C^{-1} + \lambda_1 F_Y\right]^{-1} \underline{e}_p \tag{11}$$

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

$$B\left(\underline{\theta}, \delta, \gamma\right) = \left[\underline{e}_p + \underline{d}_{\min}\right] F_Y^{-1} \left[\underline{e}_p + \underline{d}_{\min}\right] \tag{12}$$

$$\underline{d}_{\min} = -\left[F_Y^{-1} + \lambda_1 C + \lambda_2\left[M_p - \gamma^2 I\right]\right]^{-1}\left[F_Y^{-1} - \lambda_2 \gamma^2 I\right] \underline{e}_p \tag{13}$$

and $\lambda_1, \lambda_2 \ge 0$ are given by the equality constraints

$$\left(\underline{d}_{\min} + \underline{e}_p\right)^T \left[M_p - \gamma^2 I\right]\left(\underline{d}_{\min} + \underline{e}_p\right) = 0 \tag{14}$$

$$\underline{d}_{\min}{}^T C \underline{d}_{\min} - \gamma^2 = 0 \tag{15}$$

IV. Results

A. Linear Gaussian Model

We generated a delta-gamma-sigma surface for the linear Gaussian inverse problem $\underline{Y} = A\underline{\theta} + \underline{\varepsilon}$. The additive noise is distributed $N(0, \Sigma)$ with covariance $\Sigma = I$. A is a 128x128 matrix with elements $a_{ij} = (1/\sqrt{2\pi}w) \cdot \exp\left(-\frac{(i-j)^2}{2w^2}\right)$ and $w = 0.5$. The estimation task is for the 67[th] pixel (i.e. $\hat{\theta}_p(Y) = \hat{\theta}_{67}(Y)$). We also overlay on the surface the variance of the Quadratically Penalized Maximum Likelihood estimator $\hat{\underline{\theta}} = \left[F_Y + \beta P\right]^{-1} A^T \Sigma^{-1} \underline{Y}$ where $\beta \ge 0$ is a regularization parameter and $\beta \ge 0$. The penalty matrix P is the 1[st]-order neighborhood difference (Laplacian) matrix,, and the choice of norm matrix is $C = P^{-1}$.

Figure 2 shows a contour plot of this surface, along with the trace of the QPML estimator for values of the penalty $10^{-3} \le \beta \le 10^4$. Although not shown from this perspective, the variance of the QPML estimator meets the bound at all points.

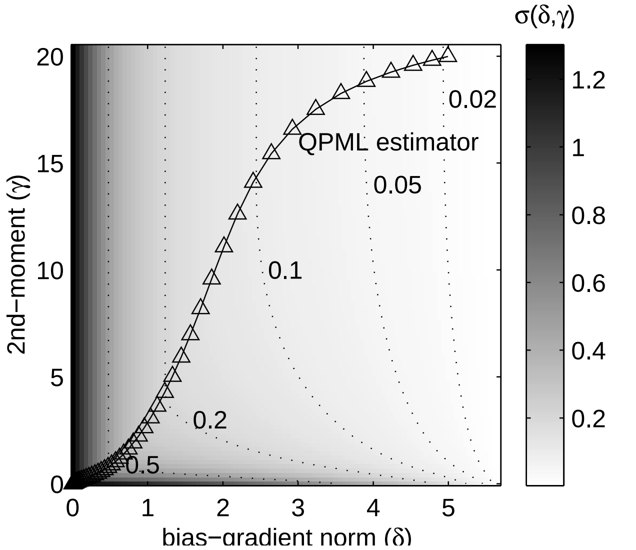

Figure 2: delta-gamma-sigma surface for linear Gaussian inverse problem. The QPML estimator is overlaid on top.

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 aperature and parallel hole collimator", Proceedings of the 1996 IEEE NSS-MIC, Anaheim CA, Nov 1996 pp. 788-792