Parameter Estimation for Multi-dimensional Filtered Poisson Processes

Alfred Hero
Dept. of EECS,
The University of Michigan,
Ann Arbor, MI 48019-2122
heroeecs.umich.edu

Abstract

As comprehensively illustrated in the classic text on point processes by Donald Snyder [1] and its revision [2] co-authored with Michael Miller, multi-dimensional Poisson processes are of great interest in many scientific and engineering applications. In many of these applications it is of interest to estimate an underlying parameter of the point process' intensity function given filtered and noise corrupted measurements. In this paper we derive lower bounds on a general class of filtered and Gaussian noise corrupted multi-dimensional Poisson processes which covers both the linear superposition "shot noise" model and the multiplicative superposition "coverage" model. The key to the lower bounding technique is an information theoretic inequality which relates the parameter estimation mean-square-error to the Shannon capacity of the measurement channel. Application of the data processing theorem to this inequality reveals SNR regimes of operation where estimator performance is limited by the quantum Poisson noise vs. by the continuous Gaussian noise. We illustrate these bounds for optical positioning and porosimetry applications.

I. Introduction

Multi-dimensional filtered Poisson processes have been used as models in many area of science and engineering such as: porosimetry and granulometry in materials science and other areas [3], [4], [5]; quantum limited photo-detection and optical-positioning using CCD arrays [6], [7], [8] or crystal scintillators [9], [10]; and electron microscopy using silver grain emulsions [11]. Frequently it is of interest to estimate parameters of these models, i.e. parameters of the point process intensity function. When the Poisson process is observed directly without error, the so-called quantum limited regime, optimal estimators of intensity parameters can easily be derived. However, practical measurement systems introduce instrumentation degradations due to measurement noise and a spatio-temporal bandlimited point spread function leading to a noise contaminated and filtered Poisson process observations. For these observations, the posterior likelihood function is not closed form, indeed it does not even have a finite dimensional representation, and estimation methods developed for the direct photon observation regime are not directly applicable. In the absence of optimal estimators it is desirable to have explicit estimator independent lower bounds on mean square error (MSE) which can be used for benchmarking, intrument design and instrument evaluation.

This paper presents a unified approach to obtaining lower bounds on MSE which covers both the linear superposition "shot noise" model and the multiplicative superposition "coverage" model for the filtered Poisson process. The key to the lower bounding technique is an information theoretic inequality which relates the parameter estimation mean-square-error to the Shannon capacity of the

[1]This work was supported in part by Air Force Office of Scientific Research, Grant F49620-99-0028.

measurement channel. Application of the data processing theorem to this inequality reveals a SNR regime of operation where estimator performance is limited by the directly observed quantum noise vs. a regime where performance is limited by the instrumentation noise. After presenting the general bound we treat two special cases: optical position estimation given spatio-temporal images of an optical beam incident on a continuous CCD array; and porosimetry given a image of a slice of porous composite material acquired by microtomography.

II. General Measurement Model

Let $\Theta = [\Theta_1, \ldots, \Theta_p]^T$ be a vector of random variables taking values $\theta = [\theta_1, \ldots, \theta_p]^T$ in $\mathbb{R}^p$ and having a joint density $f_\Theta(\theta)$. The general goal is to develop a MAP estimator of Θ and to specify lower bounds on the mean square estimation error (MSE) of its components. Estimation of Θ is based on an observed image $Y = \{Y(u) : u \in I\}$ composed of a signal image S and a noise image W. Here I will either denote a three dimensional space time domain $A \times [0, T]$, for the spatio-temporal application, or a planar patch A in the space domain, for the microtomography application. The signal S is generated by a marked point process $dM = \{dM(u) : u \in I\}$ whose distribution depends on Θ. The process dM consists of a sequence of point locations $\{U_i\}^N$ in I and a sequence of real valued marks $\{R_i\}_{i=1}^N$. Conditioned on $\Theta = \theta$ and $N = n$, $\{U_i\}_{i=1}^N$ and $\{R_i\}_{i=1}^N$ are assumed mutually independent and i.i.d. with marginal densities $f_{U|\Theta}(u|\theta)$ and $f_{R|\Theta}(r|\theta)$, respectively. Here N is a Poisson r.v. independent of Θ with rate $E[N] = \Lambda > 0$. Under these assumptions the conditional joint distribution of dM is closed form and estimation of Θ from dM is easily studied [12], [2]. This is no longer true when additive noise and blurring are introduced into the observations giving rise to a model:

$$Y(u) = S(u) + W(u), \quad u \in I \tag{1}$$

where W is assumed to be white zero mean Gaussian noise with spectral power level $N_o/2$, and $S(u)$ is

$$S(u) = h(u) \star g(u; dM). \tag{2}$$

where $\star$ denotes convolution, $h(u)$ is the point spread function of the measuring instrument, which we assume spherically symmetric, and $g(u; dM)$ is a filtered Poisson process.

Two special cases are of interest.

A. Spatio-temporal Beam Position Estimation

The problem of localizing the position on a planar detector of one or more optical point sources arises in applications such as galactic astronomy and astrometry, satellite navigation and telemetry, pulsed laser radar, optical communications, and star tracking systems for global positioning [6], [7], [8], [13]. The measurements are obtained by detecting incoherent quasimonochromatic light on an optical focal plane array of CCD photo-detectors. The measurements are noise and distortion contaminated by optical diffraction, quantum (photon) noise, and thermal (electronics) noise. Assume a far-field stationary point source generates a symmetric blur function with center of symmetry at detector position θ relative to the center $(0, 0)$ of the detector surface, here assumed to be the planar patch $A = [-a, a] \times [-a, a]$. The detector produces continuous measurements over the planar patch and over the time intrval $[0, T]$ and thus $u = (z_1, z_2, t) \in A \times [0, T]$. In this case the noiseless response of the detector can be modeled as

Linear superposition (shot noise) model:

$$g(u; dM) = \sum_{i=1}^{N} R_i p(u - U_i), \quad u \in A \times [0, T]$$

where R_i is the induced charge (called the gain sequence) deposited in the photo-detector by a given incident photon at space-time position U_i and $p(u)$ is the spatio-temporal implulse response of the photo-detector.

In optical position estimation applications it is often reasonable to assume that the random gains R_i follow a θ-indpendent distribution $f_R(r)$ while the intensity depends on θ as a spatial translation parameter. When filtered and corrupted with additive noise the above photo-detector model is a spatio-temporal extension of the one-dimensional temporal model developed in [14], [15] for timing estimation in PET imaging. This one-dimensional model has been experimentally validated for a photo-detector system consisting of a Burle 8575 photo-multiplier tube (PMT) and BGO scintillator in [16]. Our measurement model is similar to models proposed in other applications such as neurological measurement of compound action potentials, seismic measurement of reflected oil exploration probes, acoustic reverberation measurement, and photo-detection in optical fiber communications [17], [18], [19], [20]. Figure 1 shows the noiseless linear superposition response of a detector under our model as a function of the intensity of a cylindrical beam.

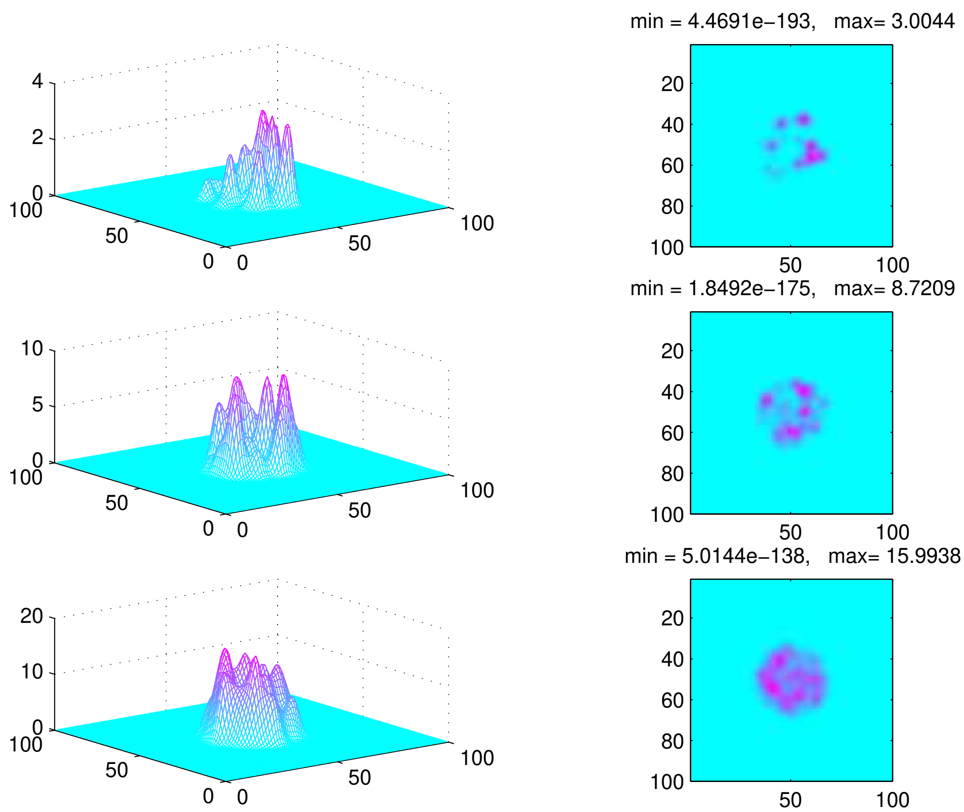

Fig. 1. *Snapshot of the spatio-temporal photo-detector (shot noise) output which is a random superposition of symmetric Gaussian-shaped single photon responses in the plane. The rows from top to bottom correspond to 50, 200 and 500 incident photons distributed uniformly over a disk shaped support.*

B. Granulometry/Porosimetry

Granulometry is the science of determining the statistical distribution of granule sizes from images of a composite material containing granules drawn from a population of grains. An application is the determination of local transparency of a photographic film composed of silver grains of random diameter [11]. Porosimetry is the science of determining the statistical distribution of pore sizes from images of a porous material. An application is the determination of local porosity from scanning electron microscope images of concrete/cement mixtures [3]. A multi-dimensional marked Poisson process model dM for these images consists of a Boolean superposition of N opaque random spheres (3D) or random disks (2D) of random radius and position. In the 2D case dM consists of disk spatial locations $U_i \in I$ where $I = A = [-a, a] \times [-a, a]$ is the image support and the marks R_i are disk radii, $i = 1, \ldots, N$. The noiseless and unfiltered 2D image $g(u, dM)$ is a set of partially occluding disks, where occlusion is caused by mutual overlapping of the opaque disks, which can be represented by the following:

Non-linear superposition (coverage) model:

$$g(u; dM) = \max_{i=1,\ldots,N} D\left(\frac{u - U_i}{R_i}\right) = \prod_{i=1}^{N} D\left(\frac{u - U_i}{R_i}\right), \quad u \in A \tag{3}$$

where $D(u)$ is the indicator function of a disc of radius 1 centered at the origin $u = (0, 0)$. Note that $g(u; dM)$ is a binary function which is non-zero only if there exists at least one disc covering the spatial point u. The Boolean model (3) is also called a "coverage process" model [21] and has been used for many applications in the life sicences, stereology, and balistics [22]. The above disk superposition model has also been used in both of the porosimetry and granulometry applications mentioned above. In these latter applications it is commonly assumed that the disk positions (spatial intensity) are uniform and independent of θ, i.e. $f_{U|\Theta}(u|\theta) = 1/|I|$, while the disk radii (mark distribution) have an unknown θ-dependent distribution $f_{R|\Theta}(r|\theta)$ which is of interest. Three representative images are shown in Fig. 2 for a linear distribution $f_{R|\Theta}(r|\theta)$ with unknown slope $\theta \in [-1, 1]$ supported on $r \in [0, 1]$ for three different values of θ.

III. Channel Decomposition and a MSE Lower Bound

Surprisingly, even though the likelihood function is intractible, at least for the linear superposition shot noise model the θ-conditional characteristic function for the filtered and noise corrupted measurement Y has a simple representation, e.g. see books by Snyder and Miller [2] or O'Reilly [20]. To contend with the intractibility of the likelihood function of Y we use a channel decomposition to derive lower bounds on the MSE of an estimator $\hat{\Theta}$.

The measurements Y are related to the parameters Θ through the conditional density $f_{Y|\Theta} = \{f_{Y|\Theta}(y|\theta)\}_{y,\theta}$ or equivalently through the log-likelihood function $l(\theta) = \ln f_{Y|\Theta}(y|\theta)$. Since Θ is a random vector of parameters we can associate $f_{Y|\Theta}$ with transition probabilities of a measurement channel C. Let X be an arbitrary random variable. Then from the Bayes identity: $f_{Y|\Theta}(y|\theta) = \int_X f_{Y|X,\Theta}(y|x, \theta) f_{X|\Theta}(x|\theta) dx$. When X satisfies $f_{Y|X,\Theta}(y|x, \theta) =$ independent of θ, the Bayes identity affirms that C is decomposable into a cascade of two channels C_1 and C_2 whose transition probabilities are, respectively, $f_{X|\Theta}$ et $f_{Y|X}$. In the language of the EM algorithm, X is a complete data set that carries more information about Θ than does Y [23]. Now, in the context of the model (1) a natural choice for X is the marked point process dM which gives the decomposition illustrated in Figure 3.

A. A Tight MSE Lower Bound via Channel Decomposition

Define the conditional mean estimator $c(dM) = E[\Theta|dM]$ of Θ given the direct observations dM and let $\hat{\Theta}(Y)$ denote an arbitrary estimator of Θ given Y. Then it is easily shown that the covariance

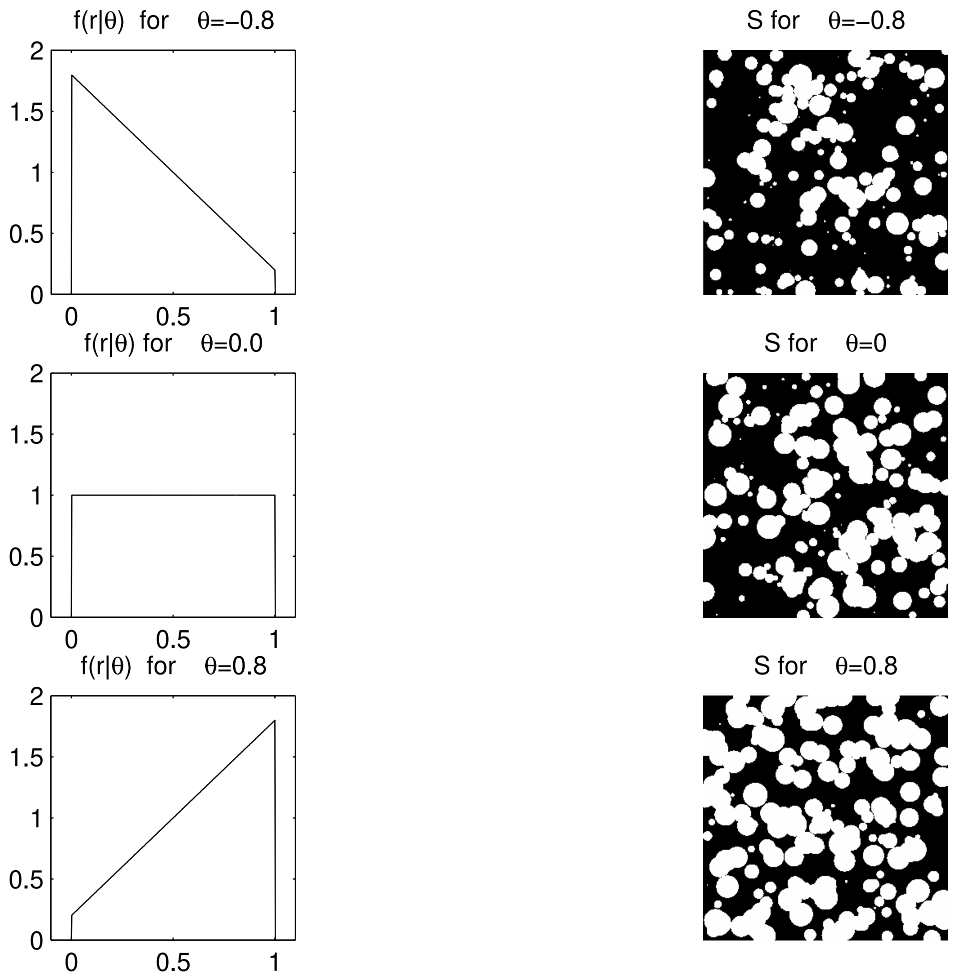

Fig. 2. *θ dependency of empirical grain distribution in a granulometry/porosimetry model. The density of disk radii is piecewise linear, with slope controlled by $\theta \in [-1, 1]$, shown at left while at right is shown a realization of the noiseless Boolean process $g(u; dM)$ for three different values of θ.*

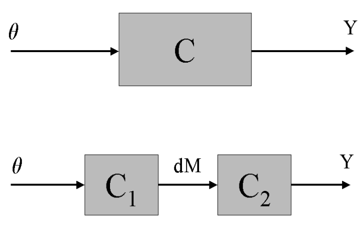

Fig. 3. *(a) Statistical representation of Y as the output of C with input Θ. (b) Decomposition of C into C_1 and C_2.*

of the estimator error satisfies the matrix bound

$$\begin{aligned} E[(\Theta-\hat{\Theta}(Y))(\Theta-\hat{\Theta}(Y))^T] &\geq E[(\Theta-E[\Theta|Y])(\Theta-E[\Theta|Y])^T] \\ &= \underbrace{E[(\Theta-g(dM)])(\Theta-g(dM)])^T]}_{\text{quantum-limited cov}} + \underbrace{E[(g(dM)-E[g(dM)|Y])(g(dM)-E[g(dM)|Y])^T]}_{\text{Gauss-limited cov}} \end{aligned} \tag{4}$$

with "=" when $\hat{\Theta}(Y))^2 = E[\Theta|Y]$. This lower bound splits the covariance into two additive components, a quantum noise limited covariance and a Gaussian noise limited covariance. As a function of SNR it can be expected that one of these two terms will dominate the other so we can talk about a quantum-limited regime of operation when the first term dominates and a Gauss-limited regime of operation when the second term dominates. An immediate consequence of the above bound is

$$E[(\Theta-\hat{\Theta}(Y))(\Theta-\hat{\Theta}(Y))^T] \geq E[(\Theta-g(dM))(\Theta-g(dM))^T]$$

which can be expected to be tight in the quantum-limited regime. Since in this regime the conditional mean can often be derived in closed form, at least in principle the lower bound can be comuted exactly. Alternatively, a number of techniques exist that can be applied to lower bound the MSE components, i.e. the diagonal elements of $E[(\Theta-g(dM)])(\Theta-g(dM)])^T]$: Schwarz-inequality bounds such as the Cramér-Rao bound (inapplicable to finite Θ however), the Bobrovsky-Zakai bound, and the Weiss-Weinstein bound; Tchebychev-inequality bounds such as the Ziv-Zakai bound and the Chernoff bound; or Shannon-inequality bounds such as the distortion-rate bound and the data processing bound. This paper deals with the latter class of bounds.

B. Distortion-Rate Bound

For an estimator $\hat{\Theta} = [\hat{\Theta}_1, \ldots, \hat{\Theta}_p]^T$ define the total mean square error $\text{MSE} = \sum_{j=1}^p E[(\Theta_i - \hat{\Theta}_i)^2]$ which is the trace of the left hand side of (4). Let V and Z be two random variables with mutual information $I(V;Z) = E[\ln P_{Z|V}(Z|V)/P_Z(Z)]$. Let $\rho(V,Z)$ be the squared distance (distortion) between the source V and an estimate $\hat{V}(Z)$ based on Z. Shannon theory asserts that for any upper bound d on $\overline{\rho} \stackrel{\text{def}}{=} E[\rho(V,Z)]$ the capacity $C = \sup_{P_V} I(V,Z)$ of the channel with input V and output Z must be at least as large as $R_\rho(d) = \inf_{P_{Z|V}:\overline{\rho}\leq d} I(V,Z)$, which is called the rate-distortion function. $R_\rho(d)$ is strictly decreasing over $d < d_{max}$ where d_{max} is the sum of the *a priori* variances of the components of V. Thus, defining the inverse $R_\rho^{-1}(\bullet)$ we have the lower bound

$$d = \text{MSE} \geq \min\{d_{max},\ R_\rho^{-1}(C)\} \tag{5}$$

To this lower bound the Shannon bound [24] can be applied: $R_\rho(d) \leq H(V) - \frac{1}{2}\ln(2\pi de)$, where $H(V) = E[-\ln P_V(V)]$. Furthermore, Shannon's data processing theorem asserts that if C is the capacity of a channel composed of a cascade of two channels with capacities C_1 and C_2 [25], then

$$C \leq \min\{C_1, C_2\}.$$

Upon application of these two Shannon bounds to (5) we obtain the following lower bound

$$\text{MSE} \geq \max\{\text{RDLB1}, \text{RDLB2}\} \tag{6}$$

where

$$\begin{aligned} \text{RDLB1} &= \frac{1}{2\pi e}\, e^{2H(V)}\, e^{-2C_1} \\ \text{RDLB2} &= \frac{1}{2\pi e}\, e^{2H(V)}\, e^{-2C_2}, \end{aligned}$$

are rate-distortion bounds which apply to the quantum-limited and Gauss-limited regimes of operation, respectively.

Now identifying the filtered Poisson process data $Y = Z$ and the parameters $\Theta = V$ in (6) and using the decomposition $C = C_1 C_2$ illustrated in Fig. 3 we can evaluate (6) once C_1 and C_2 are available. The lower bound (5) is achievable when the transition probability $P_{Y|\Theta}$ of the channel is equal to the rate distortion achieving transition probability $P_{Z|V}$ and when the prior probability P_Θ is equal to the capacity achieving prior P_V. The bound (6) is achievable when in addition, one of the capacities C_1, C_2 dominates the other.

IV. Granulometry/Porosimetry Application

For this application we assume the linear form of the conditional radii density $f_{R|\Theta}$ shown in Fig. 2 and a uniform scalar parameter Θ over the range $[-1, 1]$. In this case the CR bound is not valid since the support of f_Θ is finite. The bounds RDLB1 and RDLB2 can be obtained, however, by determining the channel capacities C_1 and C_2. These can be upper bounded by using maximum entropy arguments to produce lower bounds on RDLB1 and RDLB2.

A. Point Process Channel C_1

Using the fact that among all point processes dM with the same intensity the Poisson process has highest entropy we obtain a bound on C_1 similar to the expression obtained in [12, Lemma 4]

$$C_1 \leq C_1^* = \Lambda \sup_{f_\Theta} \int f_\Theta(\theta) \int dr\; f_{R|\Theta}(r|\theta) \ln \frac{f_{R|\Theta}(r|\theta)}{f_R(r)} d\theta$$

$$f_R(r) \stackrel{\text{def}}{=} \int f_\Theta(\theta) f_{R|\Theta}(r|\theta) d\theta,$$

C_1^* is simply the capacity of a purely Poisson channel which is equal to the maximum mean Kullback distance between the conditional density $f_{R|\Theta}(r|\theta)$ and the marginal $f_R(r)$. Thus $C_1 = 0$ when $f_{R|\Theta}(r|\theta)$ is constant in θ and is consequently identical to $f_R(r)$. In this case neither dM nor Y carry any information about Θ. For the case of a linear radial density $f_{R|\Theta}(r|\theta)$, $\theta \in [-1, 1]$, the source density f_Θ which attains capacity C_1^* is easily determined and has an approximately quadratic form, as indicated in Figure 4. The resulting capacity is the linearly increasing function of Λ: $C_1^* = \Lambda a$ where $a \approx 0.0698$.

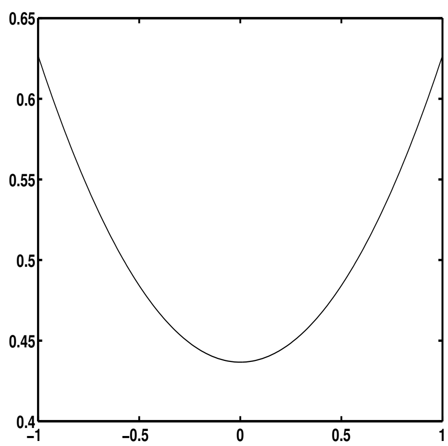

Fig. 4. *The density f_Θ^* that maximises the mutual information $I(\Theta, dM)$ and attains capacity for the case of linear $f_{R|\Theta}$ as shown in Figure 2.*

B. Continuous Process Channel C_2

Using the fact that among all continuous processes Y with fixed covariance function the Gaussian process has highest entropy we obtain the following bound under the assumption of large detector area $|A| \gg 1$ [26]

$$C_2 \leq C_2^* = \frac{|I|}{2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \ln\left(1 + \frac{\Phi_S(\omega)}{N_o/2}\right) d\omega \tag{7}$$

where $\Phi_S(\omega)$ is the power spectral density of the signal component S.

Define the function $p(u) = (1 - \|u\|/2)^{2+\lambda}$ and its Fourier transform $P(\|\omega\|)$. By making a rectangular to polar coordinate transformation in (7) we obtain the simplification

$$C_2^* = \pi |I| \int_0^{\infty} \rho \ln\left(1 + \kappa \, \frac{|H(\rho)|^2 P(\rho)}{N_o/2}\right) d\rho \tag{8}$$

where, $M_\theta(t) = E[e^{t\theta}]$ is the characteristic function of f_Θ and

$$\kappa = e^{-\pi\lambda/3}\left[M_\theta(-\pi\lambda/3) - M_\theta^2(-\pi\lambda/6) e^{-\pi\lambda/3}\right].$$

V. Beam Position Estimation Application

For this application we assume the following. The position Θ of the beam on the planar detector is uniformly distributed over the detector surface A. Conditioned on $\Theta = \theta$ the space-time Poisson process has an intensity of Gaussian form

$$\lambda(z,t) = \frac{\Lambda}{2\pi\sigma_\lambda^2} e^{-\frac{\|z-\theta\|^2}{2\sigma_\lambda^2}}, \qquad (z,t) \in A \times [0,T]$$

$$p(z,t) = \frac{1}{2\pi\sigma_p^2} e^{-\frac{\|z\|^2}{2\sigma_p^2}} \frac{1}{T_p} e^{-t/T_p} \qquad (z,t) \in A \times [0,T]$$

and the random gain sequence R_i has Gaussian density

$$f_R(r) = \mathcal{N}(\mu_R, \sigma_R^2)$$

Then, under the assumption $\sigma_\lambda^2 \ll |A|$, it can be shown [10] using analogous arguments as were used above to derive the upper bound C_1^* for the granulometry/porosimetry example

$$C_1^* = \Lambda \ln\left(\frac{|A|}{2\pi e \sigma_\lambda^2}\right)$$

and, under the assumptions that $\sigma_R^2 \gg \mu_R^2$, $T_p \ll T$, $\sigma_p^2 \ll |A|$,

$$C_2^* = \frac{|A|}{4\pi T_p \sigma_p^2}\left(\sqrt{1+\beta} + \tfrac{1}{2}\ln\left[\frac{4(\sqrt{1+\beta}-1)}{\beta(\sqrt{1+\beta}+1)}\right]\right)$$

where

$$\beta = \Lambda\left(\frac{\mu_R^2 + \sigma_R^2}{N_0/2}\right)$$

Furthermore, the direct detection (quantum-limited) CR Bound can be derived under the large detector size assumption: $\text{MSE} \geq \text{CRB1} = \sigma_\lambda^2/\Lambda$

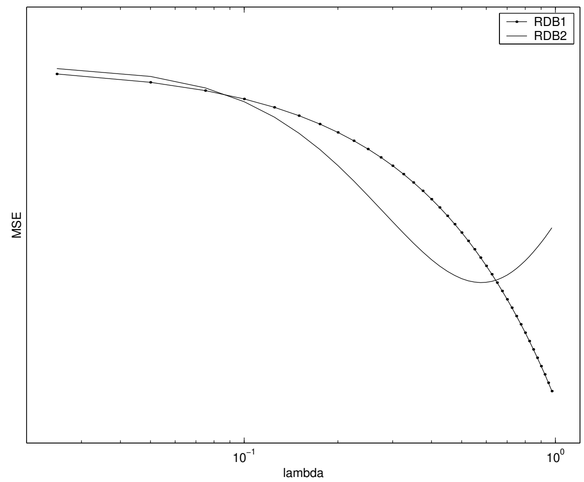

Fig. 5. *Bounds as function of $\lambda = \Lambda/|I|$ for granulometry/porosimetry example (range of MSE axis is 0.2 to 0.24).*

VI. Results

We first consider the granulometry/porosimetry example. In Figure 5, the pair of quantum-limited and Gauss-limited rate-distortion-bounds (labeled RDB1 and RDB2) are plotted as a function of the uniform spatial intensity $\lambda = \Lambda/|A|$ for uniform density $f_\Theta(\theta)$ over $[-1, 1]$ and the same values for A and σ_p as used in Figure 2. The SNR is equal to 0dB. Recall that it is the maximum of these two bounds that determine the Shannon bound (6). Notice that RDLB1 decreases monotonically in λ: estimation Θ from direct measurements dM always benefits from an increase in the number of points N. On the other hand, RDLB2 takes a minimum value and subsequently increases as λ becomes large: estimates of the disk radii distribution parameter Θ based on degraded measurements Y suffers from an increasing number of occlusions that must occur as the number of disks become large.

The bound illustrated in Fig. 5 thus separates estimator performance into two operating regions: the quantum-limited region $RDLB1 > RDLB2$ and the Gauss-limited region $RDLB2 > RDLB1$. In particular there are three distinct regions $\lambda \in [0, 0.1]$, $\lambda \in (0.1, 0.65]$ and $\lambda > 0.65$ where dominance occurs. The only region where quantum-limited operation is attainable from measurements Y is for moderate values of λ: $\lambda \in (0.1, 0.65]$. In general the boundaries of the λ regions depend on detector area $|A|$, noise level N_o, point spread function σ and prior f_Θ.

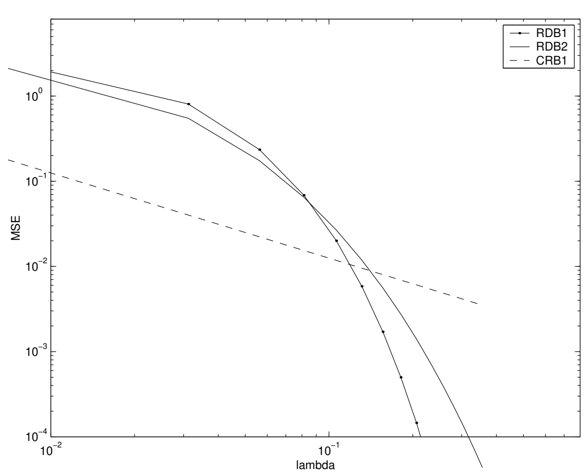

Fig. 6. *Bounds as function of $\lambda = \Lambda/|I|$ for beam positioning example. The Greatest lower bound is given by the envelope tracing out the maximum of the two curves.*

The degradation in performance for large λ is to be contrasted with the case of the position estimation example shown in Fig. 6. In this figure three bounds RDLB1, RDLB2 and the CR bound (CRB1) are plotted. For the choice of operating parameters σ_p, T_p, N_o chosen there is a region of λ where one of each bound dominates the others. The decrease in achievable MSE occurs in this example since, as there is no mutual occlusion of the superimposed impluse responses, the SNR monotonically increases as λ increases.

VII. Conclusions

We have derived useful lower bounds on MSE for estimation of random parameters of the intensity of filtered multidimensional conditionally Poisson processes. These bounds were obtained using information theoretic inequalities. We established that for the position estimation problem achievable accuracy improves as photon count rate increases. On the other hand, for the granulometry problem achievable accuracy is not monotone in particle count rate. The bounds are tight for low SNR for which the minimum achievable MSE is close to the a priori variance of the parameter. Approximations to the MAP estimator have been derived elsewhere for the granulometry example [26] and for the beam position estimation example [10]. We are currently performing comparisons between the experimental MSE of these estimators and the lower bounds presented here.

References

[1] D. L. Snyder, *Random Point Processes*, Wiley, New York, 1975.

[2] D. L. Snyder and M. I. Miller, *Random Point Processes in Time and Space*, Springer-Verlag, New York, 1991.

[3] E. J. Garboczi, "Mercury porosimetry and effective networks for permeability calculations in porous materials," *Powder Technology*, vol. 67, no. 121, , 1991.

[4] D. P. Bentz, E. J. Garboczi, C. J. Haecker, and O. M. Jensen, "Effects of cement particle size distribution on performance properties of portland cement-based materials," *Cement and Concrete Research*, vol. 29, no. 10, pp. 1663–1671, 1999.

[5] B. D. Ripley, *Spatial statistics*, Wiley, New York, 1981.

[6] J. Amoss and F. Davidson, "Detection of weak optical images with photon counting techniques," *Applied Optics*, vol. 11, pp. 1793–1799, 1972.

[7] L. Kazovsky, "Beam position estimation by means of detector arrays," *Opt. Quantum Electron.*, vol. 13, pp. 201–208, 1981.

[8] P. Salomon and T. Glavich, "Image signal processing in sub-pixel accuracy star trackers," in *Proc. Soc. Phot-Opt. Instrum. Eng.*, pp. 64–74, 1980.

[9] N. H. Clinthorne, W. L. Rogers, L. Shao, and K. Koral, "A hybrid maximum likelihood position computer for scintillation cameras," *IEEE Trans. Nuclear Science*, vol. NS-37, no. 2, pp. 658–663, 1990.

[10] A. O. Hero, "Recovering photon-intensity information from continuous photo-detector measurements," in *Proceedings of the 25-th Conference on Information Sciences and Systems*, pp. 643–648, Johns Hopkins University, Baltimore, MD, Mar. 1991.

[11] B. Picinbono, "Modèle statistique suggéré par la distribution de grains d'argent dans les films photographiques," *Comptes Rendus de l'Académie des Sciences*, vol. Séance du 6 Juin, pp. 2206–2208, 1955.

[12] A. O. Hero, "Lower bounds on estimator performance for energy invariant parameters of multi-dimensional Poisson processes," *IEEE Trans. on Inform. Theory*, vol. 35, pp. 843–858, July 1989.

[13] K. Winick, "Cramer-Rao lower bounds on the performance of charge-coupled-device optical position estimators," *J. Opt. Soc. Am.*, vol. 3, pp. 1809–1815, Nov. 1986.

[14] A. O. Hero, N. Antoniadis, N. H. Clinthorne, W. L. Rogers, and G. D. Hutchins, "Optimal and sub-optimal post-detection timing estimators for PET," *IEEE Trans. Nuclear Science*, vol. NS-37, no. 2, pp. 725–729, 1990.

[15] A. O. Hero, "Timing estimation for a filtered Poisson process in Gaussian noise," *IEEE Trans. on Inform. Theory*, vol. 37, no. 1, pp. 92–106, Jan. 1991.

[16] N. A. Petrick, A. O. Hero, N. H. Clinthorne, and W. L. Rogers, "A method for improved time-of-arrival estimation for weak optical pulses with applications to scintillation detectors," *IEEE Trans. Nuclear Science*, vol. NS-38, no. 2, pp. 174–177, 1991.

[17] A. M. Bruckstein, T. J. Shan, and T. Kailath, "The resolution of overlapping echos," *IEEE Trans. Acoust., Speech, and Sig. Proc.*, vol. ASSP-33, pp. 1357–1368, Dec. 1985.

[18] J. M. Mendel, "White noise estimators for seismic data processing in oil exploration," *IEEE Trans. Automatic Control*, vol. AC-22, no. 5, pp. 694–706, Oct. 1977.

[19] P. Faure, "Theoretical models of reverberation noise," *J. Acoust. Soc. Am.*, vol. 36, pp. 259–268, 1964.

[20] J. J. O'Reilly, "Generating functions and bounds in optical communications," in *Problems of randomness in communications engineering*, K. Cattermole and J. O'Reilly, editors, chapter 7, pp. 119–133, Wiley, New York, Nov. 1987.

[21] P. Hall, *Introduction to the theory of coverage processes*, Wiley, New York, 1988.

[22] H. Elias, *Stereology*, Springer, Berlin, 1967.

[23] A. O. Hero and J. A. Fessler, "A recursive algorithm for computing CR-type bounds on estimator covariance," *IEEE Trans. on Inform. Theory*, vol. 40, pp. 1205–1210, July 1994.

[24] T. Berger, *Rate Distortion Theory: A Mathematical Basis for Data Compression*, Prentice-Hall, Englewood Cliffs NJ, 1971.

[25] R. G. Gallager, *Information Theory and Reliable Communication*, Wiley, 1968.

[26] A. O. Hero, "Sur un problème d'estimation pour des processus de poisson composeés et filtrés," *Traitement du Signal*, vol. 15, no. 6, pp. 493–502, 1999.