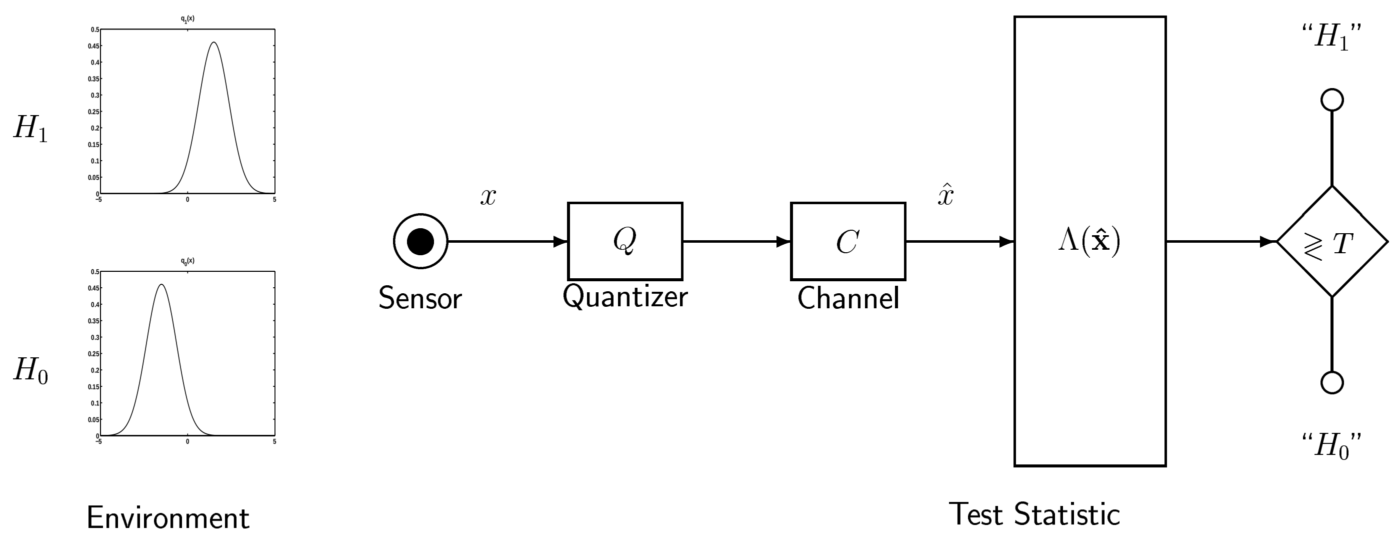
PERFORMANCE LIMITS OF HYPOTHESIS TESTING FROM VECTOR-QUANTIZED DATA

Riten Gupta and Alfred O. Hero, III
Dept. of Electrical Engineering and Computer Science
University of Michigan

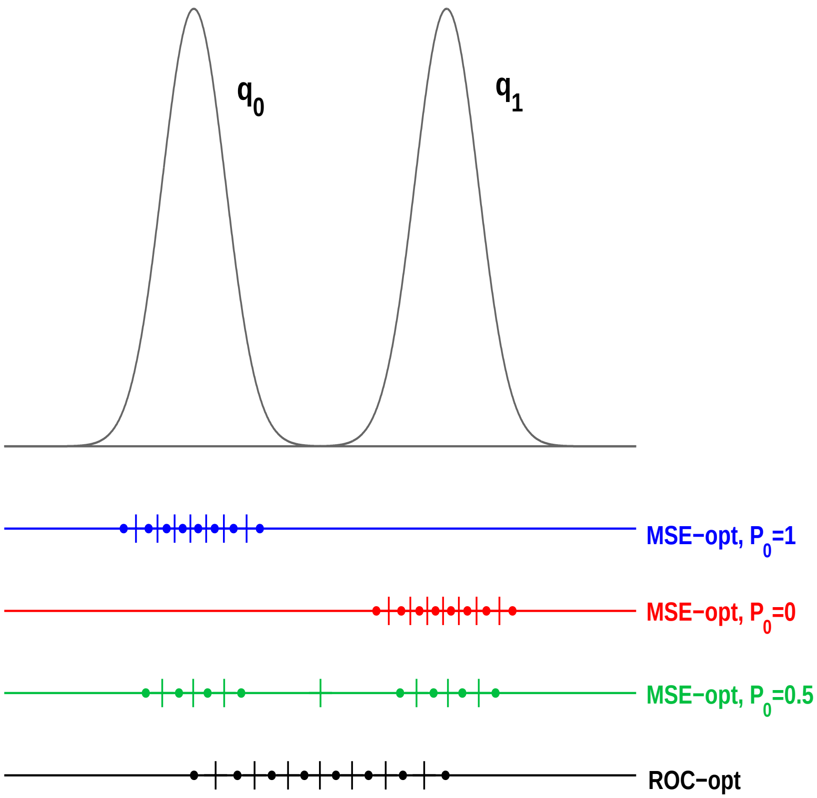
June , 2001

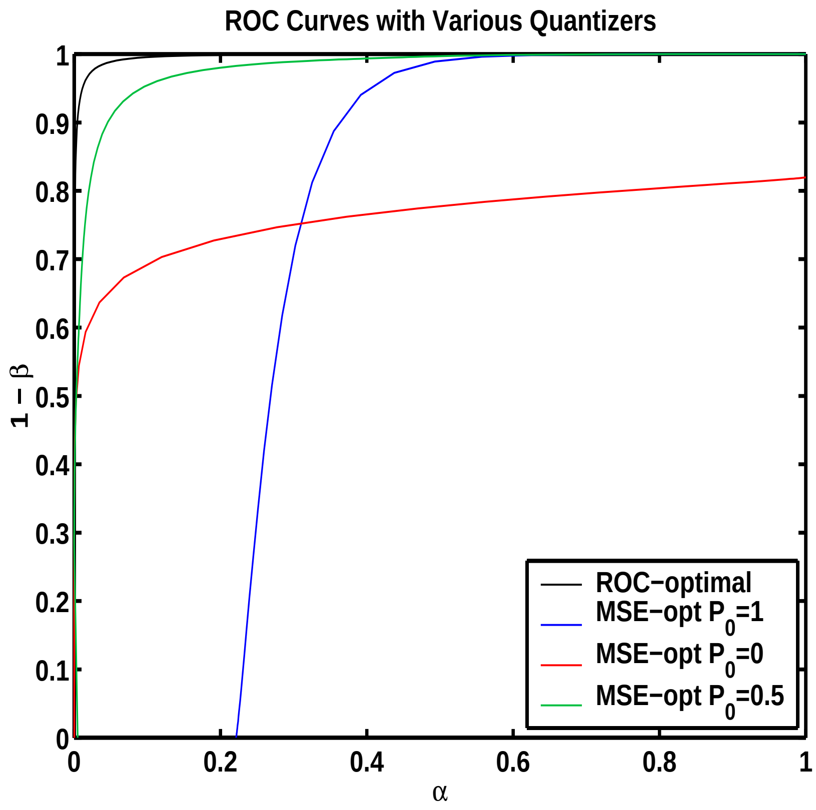
Hypothesis testing from Q/VQ data

Objective: Observe $\hat{x} = Q(x)$ and decide between H_0 and H_1



ROC Performance of Various Quantizers





Likelihood Ratio Test

Given $\mathbf{x} = [x^{(1)}, \ldots, x^{(n)}]^T$, optimal test of $H_0 \;:\; x \sim q_0$ vs $H_1 \;:\; x \sim q_1$

$$\mathbf{\Lambda}(\mathbf{x}) = \log \frac{q_1(\mathbf{x})}{q_0(\mathbf{x})} \;\underset{H_0}{\overset{H_1}{\gtrless}}\; T$$

- False alarm probability $P_F(T) = \alpha$

$$P_F(T) = P\left(\mathbf{\Lambda}(\mathbf{x}) > T \,|H_0\right)$$

- Probability of miss: $P_M(T) = \beta$

$$P_M(T) = P\left(\mathbf{\Lambda}(\mathbf{x}) < T \,|H_1\right)$$

- Probability of Detection (power): $P_D(T) = 1 - \beta$

Existence of Lossless Quantizers for Detection

- H_1 decision region for threshold T

$$\mathcal{X}_1(T) = \{\mathbf{x} \,:\, \mathbf{\Lambda}(\mathbf{x}) \geq T\}$$

- $\mathbf{Q}$ is **sufficient quantizer** if it preserves $\mathcal{X}_1(T)$ for *all* T and α

  $\Leftrightarrow \mathbf{Q}(\mathbf{x})$ is a sufficient statistic

- $\mathbf{Q}$ is a **Neyman-Pearson quantizer of level** α if it preserves $\mathcal{X}_1(T)$ for

$$\{T \,:\, P_F(T) = \alpha\}$$

Example of a Sufficient Quantizer

Figure 1:

Sufficient quantizer for 1-D piecewise-constant sources.

NP Q has Poor MSRE

Quantized Log-Likelihood Ratio (LLR) is a NP Quantizer:

$$x \longrightarrow \boxed{\Lambda(x)} \longrightarrow \boxed{Q_\Lambda} \longrightarrow$$

Log-likelihood ratio Scalar quantizer

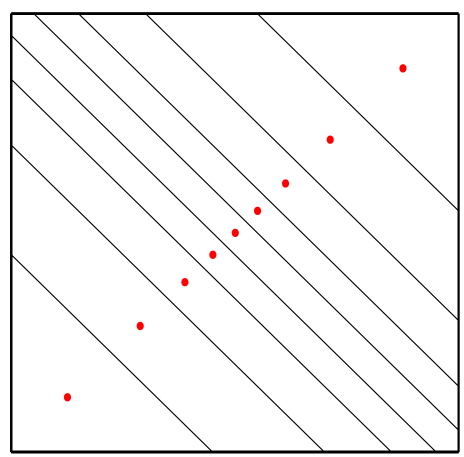


Figure 2: LR VQ when sufficient statistic is $x^{(1)} + x^{(2)}$

Problem: *Good detection* performance, but *poor estimation* performance

Detection-optimal Q/VQ with small MSRE

1. Constrained detection design objective
   $\max_Q\{J_1\}$ subject to MSRE $\leq \delta$
   where
   $$J_1 = \text{Post Q Decision Error}$$
2. Mixed detection-estimation design objective ($\rho \in [0,1]$)
   $\max_Q\{J_2\}$ where
   $$J_2 = (1-\rho) \cdot \text{Post Q Decision Error} + \rho \cdot \text{Reconstruction MSE},$$

Some post-Q detection error criteria

:

1. Bayes risk (Oehler, Gray 95, Pearlmutter *etal* 96)

$$P_e = P_M\ P(H_1) + P_F\ P(H_0)$$

2. KL and Chernoff Information (Poor 77, 78; Benitz, Bucklew 89; Jana, Moulin , Ramchandran 99)

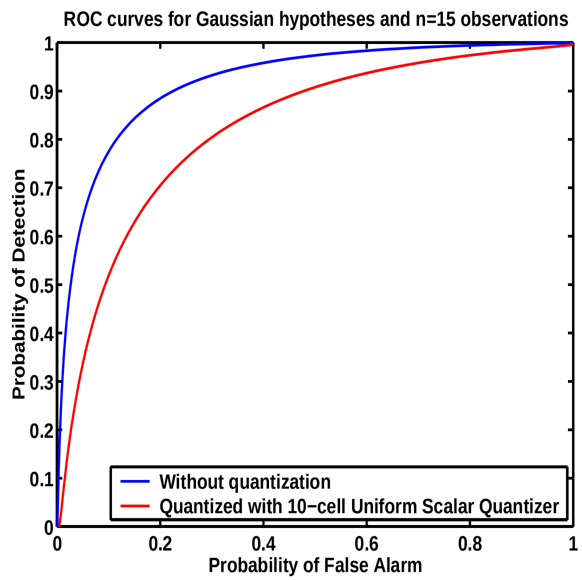
$$L = n^{-1}\log P_e$$

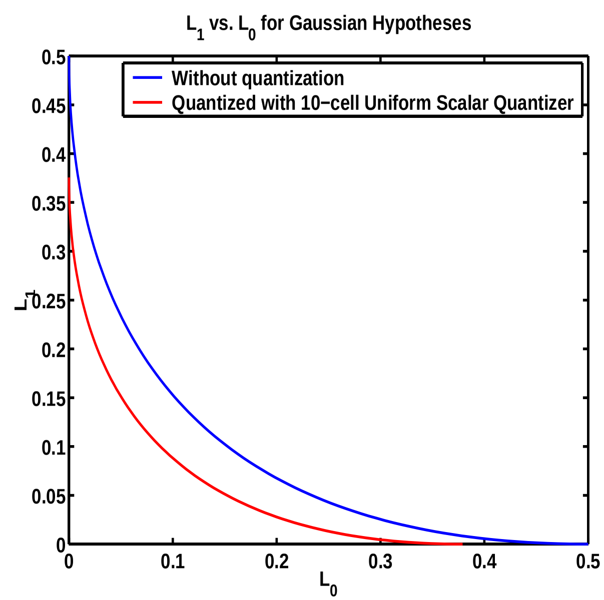
3. Sanov Information (Gupta, Hero 99)

$$L_0 = n^{-1}\log P_F, \qquad L_1 = n^{-1}\log P_M$$

4. SNR (Picinbono, Duvaut 85; Tsitsiklis 93)

Area Under the Curve (AUC) Criterion





$$\mathrm{AUC}_1 = \int_0^1 (1-\beta)\, d\alpha, \qquad \mathrm{AUC}_2 = \int_0^1 L_1\, dL_0$$

Advantage of AUC: captures detection error independent of threshold T

Special Case

For detection of shift in μ in $\mathcal{N}(\mu, \sigma^2)$

$$\mathrm{AUC}_1 = \frac{1}{2} + \frac{1}{2}\mathrm{Erf}\left(\frac{\sqrt{SNR}}{2}\right)$$

$$\mathrm{AUC}_2 = cSNR$$

Large Deviations Error Exponents for LRT

Sanov's theorem: for n large:

$$\begin{aligned}\alpha &\approx e^{-nL(q_\lambda \| q_0)}\\ \beta &\approx e^{-nL(q_\lambda \| q_1)}.\end{aligned}$$

Where, KL distance is

$$L(q_1, q_0) = \int q_0(x) \ln \frac{q_0(x)}{q_1(x)} dx$$

and for $\lambda = f(T)$, $\lambda \in [0, 1]$:

$$q_\lambda(x) = \frac{q_0(x)^{1-\lambda} q_1(x)^\lambda}{\int q_0(y)^{1-\lambda} q_1(y)^\lambda dy} = \text{``tilted''} \text{ density}$$

Note:

- λ determines T and level α of LRT
- λ for minimax LRT satisfies:

$$L(q_\lambda \| q_0) = L(q_\lambda \| q_1)$$

- As λ parameterizes curve (L_0, L_1)

$$\mathrm{AUC} = \int_0^1 L_1 \, dL_0 = \int_0^1 L_1(\lambda) \frac{dL_0(\lambda)}{d\lambda} \, d\lambda$$

For Q cells $\{S_i\}_{i=1}^N$ define pmf's of Quantized $\mathbf{x}$

$$\bar{q}_0(i) = P(x \in \mathcal{S}_i \mid H_0), \qquad \bar{q}_1(i) = P(x \in \mathcal{S}_i \mid H_1)$$

High-Resolution Analysis

Define distortions for a $\log_2 N$ bit Q

$$\Delta L_{0,N} \stackrel{\text{def}}{=} L(\bar{q}_\lambda \| \bar{q}_0) - L(q_\lambda \| q_0)$$

$$\Delta L_{1,N} \stackrel{\text{def}}{=} L(\bar{q}_\lambda \| \bar{q}_1) - L(q_\lambda \| q_1)$$

High-resolution representation:

$$\Delta L_{j,N} = N^{-2/k} \left(\lim_{N \to \infty} N^{2/k} \Delta L_{j,N} \right) + o(N^{-2/k})$$

Q is **optimal high-rate** if high-resolution distortion = min

Functions Associated with High-Rate Q

Specific point density function of cell positions (Na&Neuhoff 95):

$$\zeta_s(x) = \frac{1}{NV_i}, \text{ for } x \in S_i,$$

Specific inertial profile of cell shape (Na&Neuhoff 95):

$$m_s(x) = \frac{\int_{S_i} \|y - x_i\|^2 dy}{V_i^{1+2/k}}, \text{ for } x \in S_i,$$

Specific **covariation profile** of cell shape:

$$M_s(x) = \frac{\int_{S_i} (y - x_i)(y - x_i)^T dy}{V_i^{1+2/k}}, \text{ for } x \in S_i.$$

Divergence: Asymptotic Forms

$$\Delta L_{0,N} \approx \frac{1}{2N^{2/k}} \int \frac{q_\lambda(x)\mathcal{F}(x)}{\zeta(x)^{2/k}} \left[\lambda^2 + \lambda(1-\lambda)(L(q_\lambda \| q_0) - \Lambda_0(x))\right] dx$$

$$\Delta L_{1,N} \approx \frac{1}{2N^{2/k}} \int \frac{q_\lambda(x)\mathcal{F}(x)}{\zeta(x)^{2/k}} \left[(1-\lambda)^2 + \lambda(1-\lambda)(L(q_\lambda \| q_1) - \Lambda_1(x))\right] dx$$

where

$$\mathcal{F}(x) = \nabla\Lambda(x)^T M(x) \nabla\Lambda(x) = \text{Fisher covariation profile}$$

$$\Lambda_0(x) = \log \frac{q_{\lambda(x)}}{q_0(x)} \qquad \Lambda_1(x) = \log \frac{q_{\lambda(x)}}{q_1(x)}.$$

AUC-Optimal High-Resolution Q/VQ

Loss in area under (L_0, L_1) curve:

$$\Delta A_N \approx \frac{1}{2N^{2/k}} \int \frac{\mathcal{F}(x)\eta(x)}{\zeta(x)^{2/k}} dx.$$

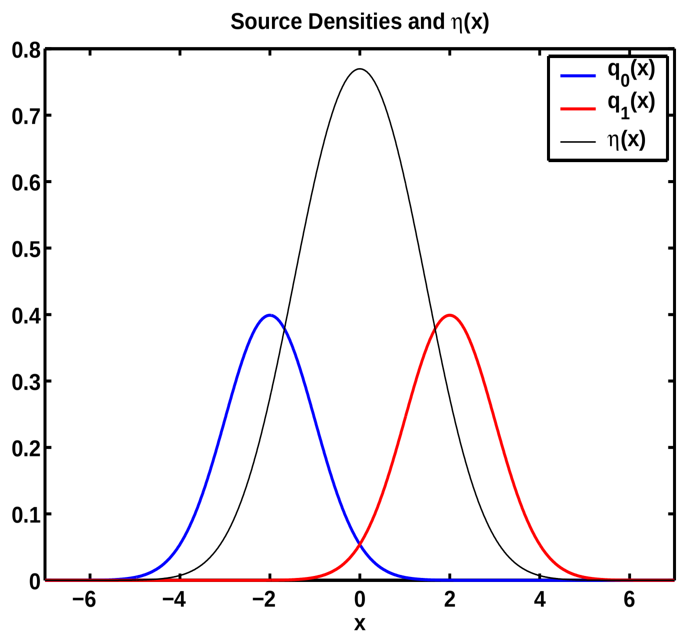
Optimal point density:

$$\zeta^o(x) = \frac{[\mathcal{F}(x)\eta(x)]^{\frac{k}{k+2}}}{\int [\mathcal{F}(y)\eta(y)]^{\frac{k}{k+2}} dy}.$$

$\eta(x) = \int_0^\infty \Psi(\lambda)\, q_\lambda\, d\lambda$: AUC-mean tilted density

1-D Gaussian Example: Source Densities and Point Densities

$$q_0 \sim \mathcal{N}(-2, 1), \qquad q_1 \sim \mathcal{N}(2, 1), \qquad k = 1$$



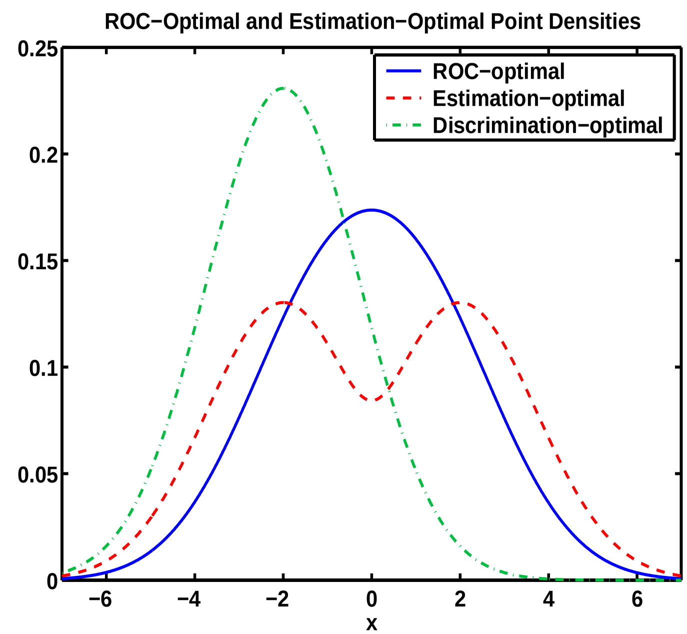
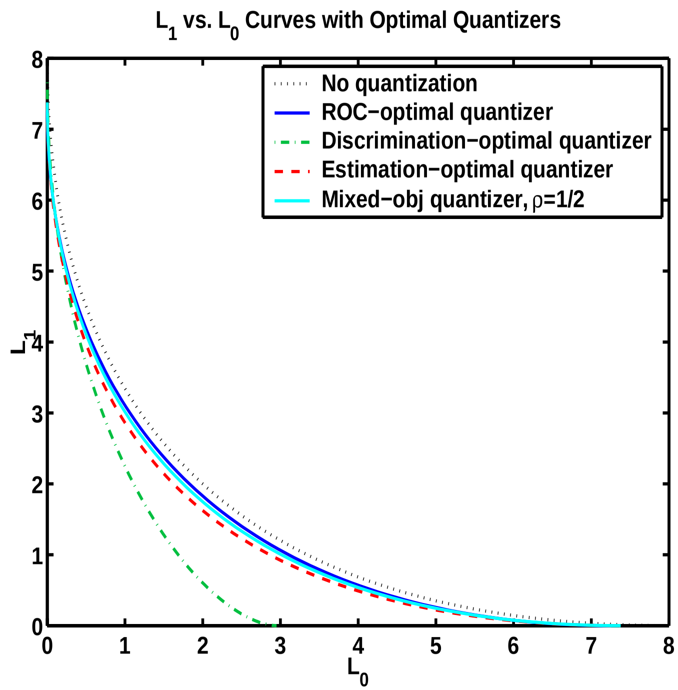


Figure 3: Source densities (left) and optimal point densities (right).

1-D Gaussian Example: Detection Performance



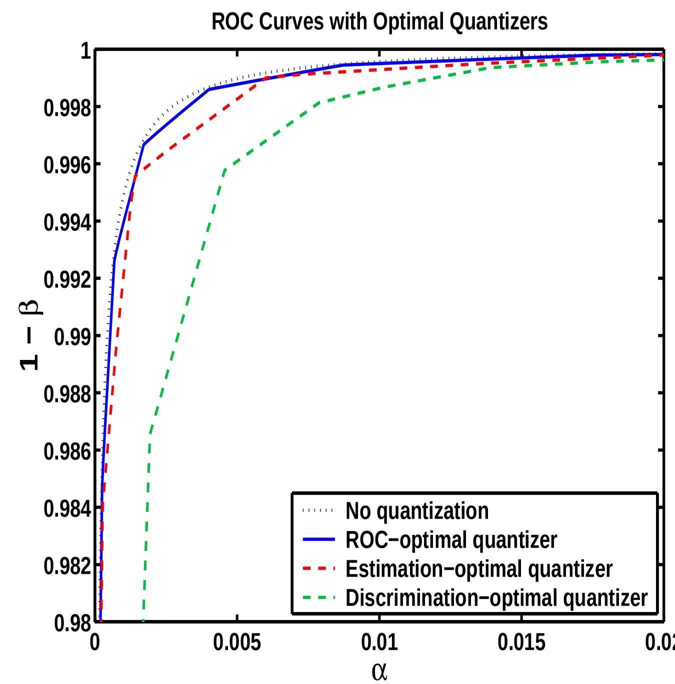


Figure 4: $L_1(L_0)$ curves with $N = 8$ (left) and ROC curves with $N = 16$ and $n = 2$ (right).

$\alpha = 0.2$ for discrimination-optimal VQ

1-D Gaussian Example: Minimax point density

$$q_0 \sim \mathcal{N}(-4, 1), \qquad q_1 \sim \mathcal{N}(4, 1), \qquad k = 1$$

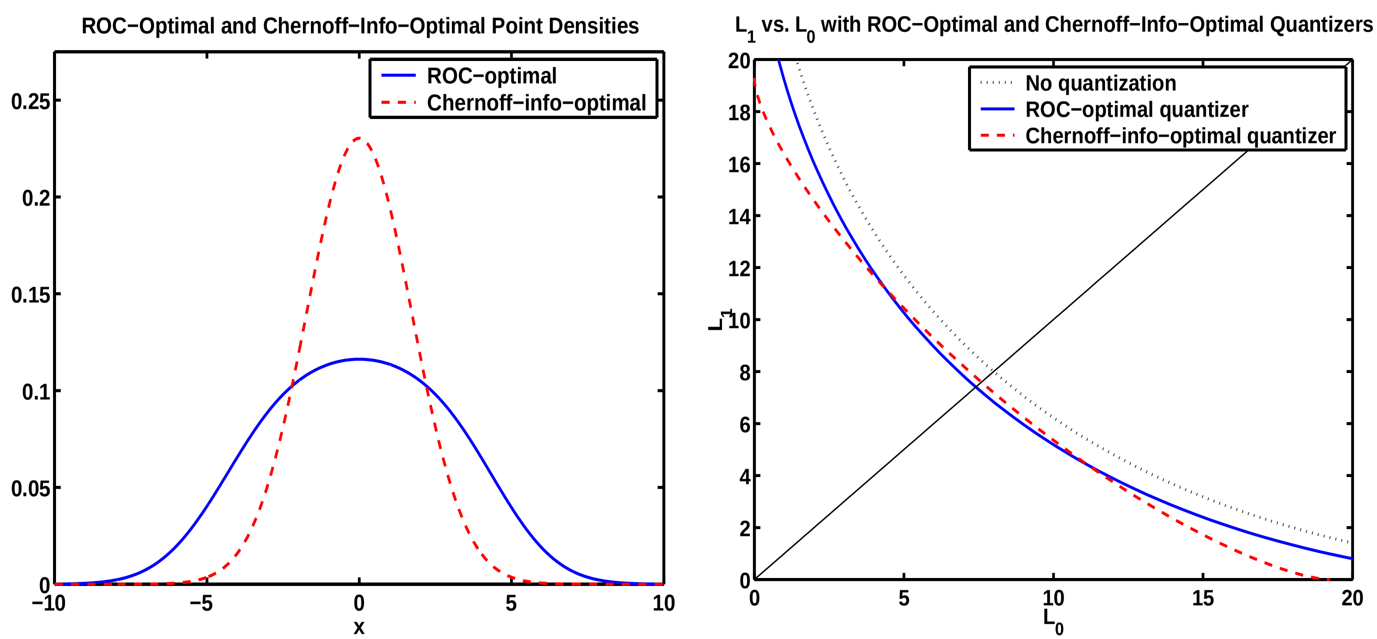


Figure 5: ROC-optimal and Chernoff-information-optimal point densities (left) and $L_1(L_0)$ curves with ROC-optimal and Chernoff-information-optimal quantizers with $N = 8$ (right).

1-D Gaussian Example: Estimation Performance (MSRE)

Figure 6: *MSRE with ROC-optimal, detection-optimal, and estimation-optimal VQ's with* $N = 16$.

2-D Anisotropic Gaussian Example

Figure 7: *Source densities for 2-D anisotropic Gaussian example.*

2-D Anisotropic Gaussian Example

Figure 8: *Two-dimensional anisotropic Gaussian example: (a)* $\eta(x)$, *(b) log-likelihood ratio* $\Lambda(x)$, *(c) discriminability* $\|\nabla\Lambda(x)\|^2$, *(d) ROC-optimal point density, (e) discrimination-optimal point density, (f) estimation-optimal point density.*

2-D Anisotropic Gaussian Example

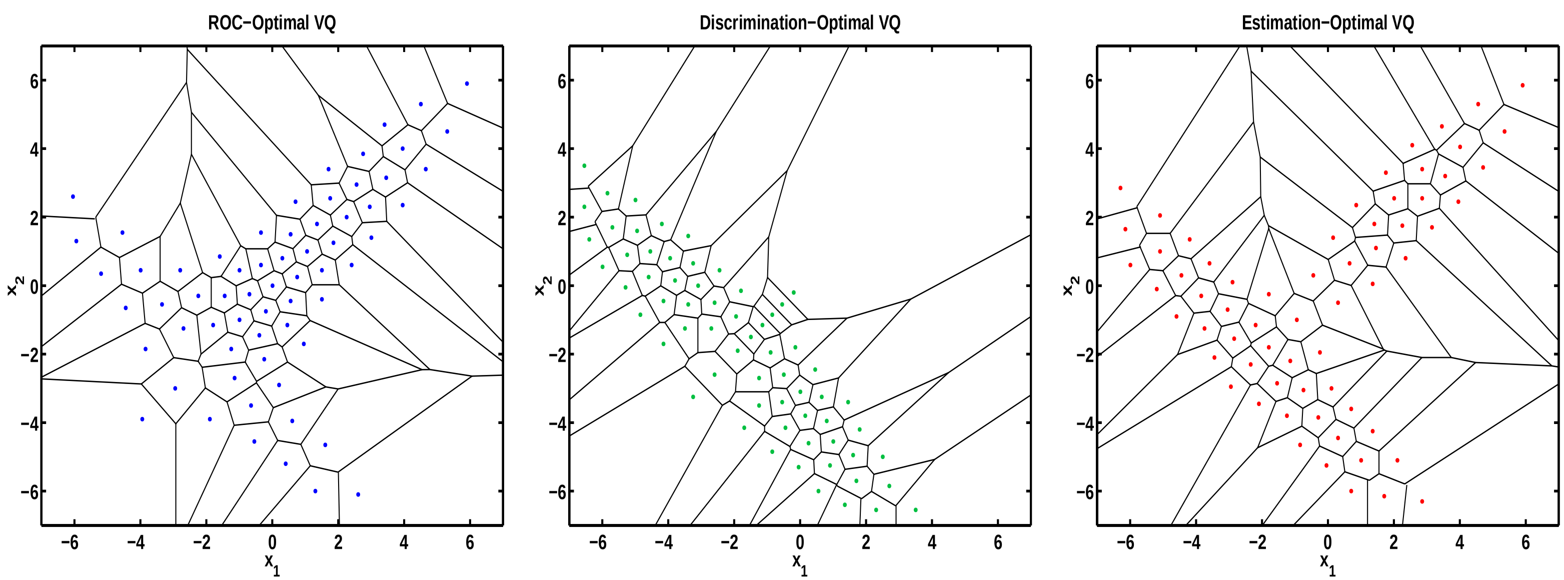


Figure 9: ROC-optimal congruent-cell VQ (left), Discrimination-optimal congruent-cell VQ (middle), and Estimation-optimal VQ (right) with $N = 64$.

2-D Anisotropic Gaussian Example

Figure 10: $L_0, L_1)$ *curves for ROC-optimal, estimation-optimal, and discrimination-optimal congruent-cell VQ's with* $N = 64$.

Medical imaging Application

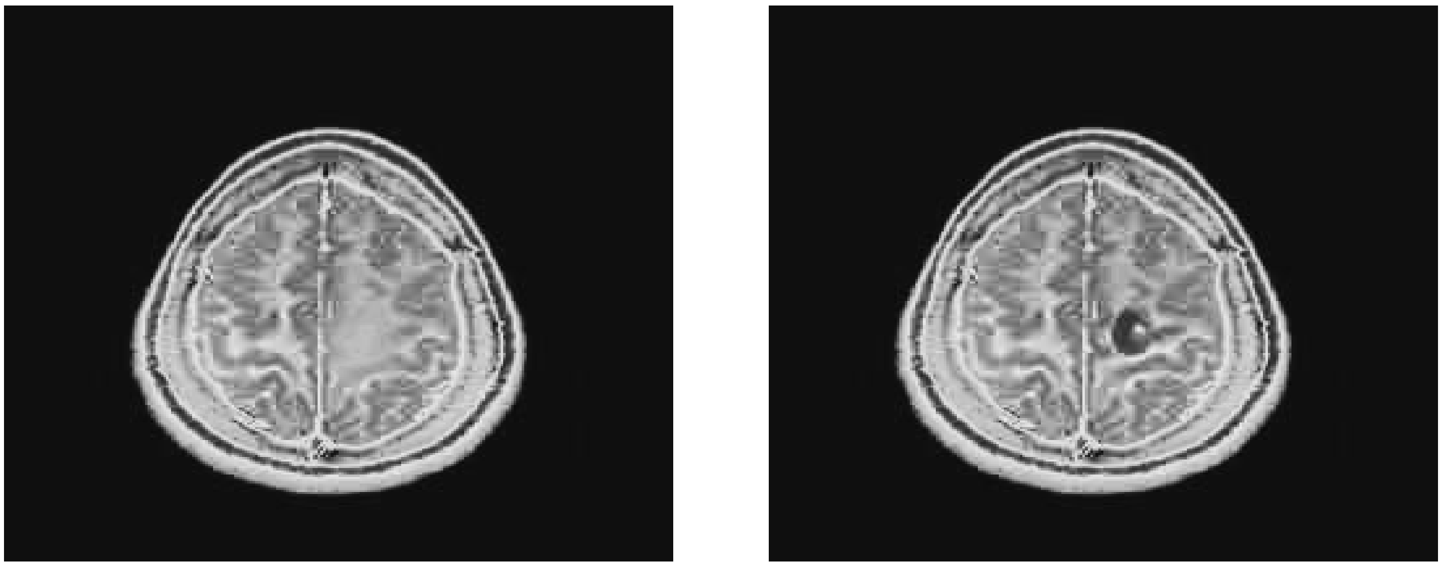
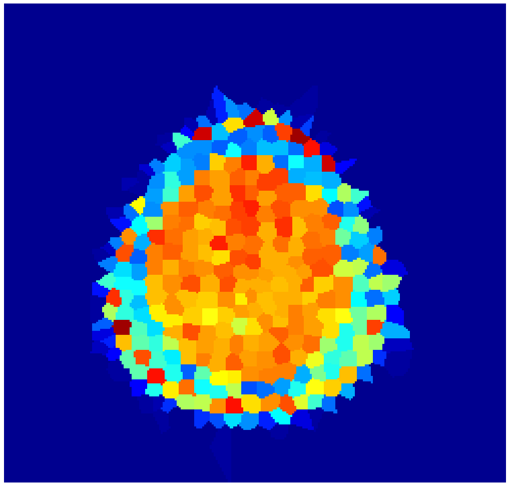
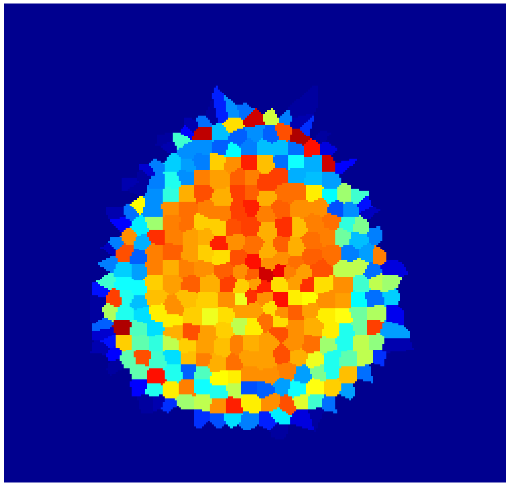


Figure 11: Pre-operative (left), Post-operative (right).

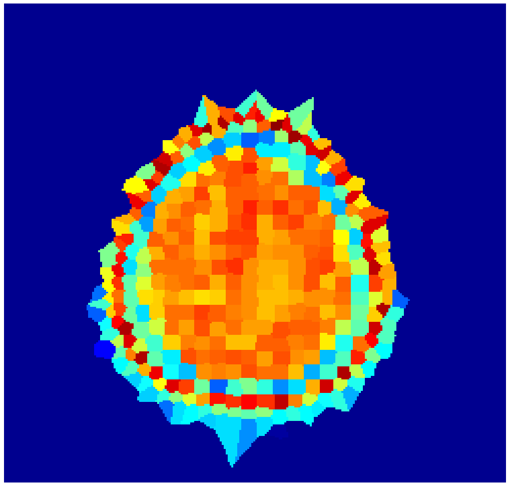
H0 with Detection VQ



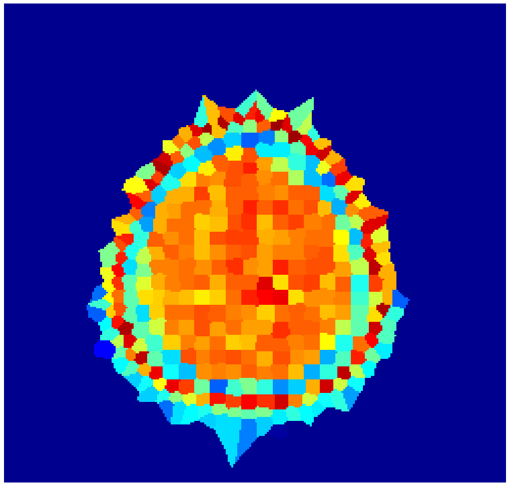
H1 with Detection VQ



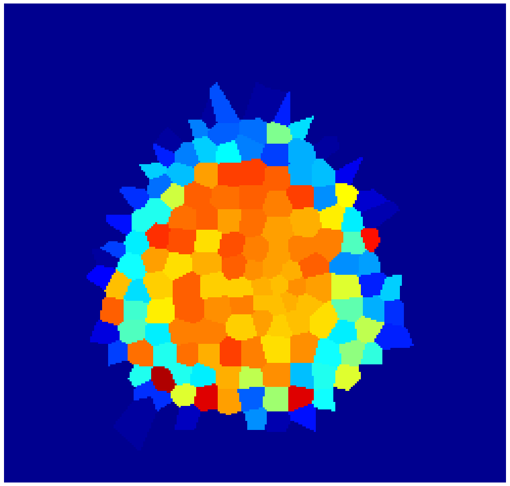
H0 with Estimation VQ



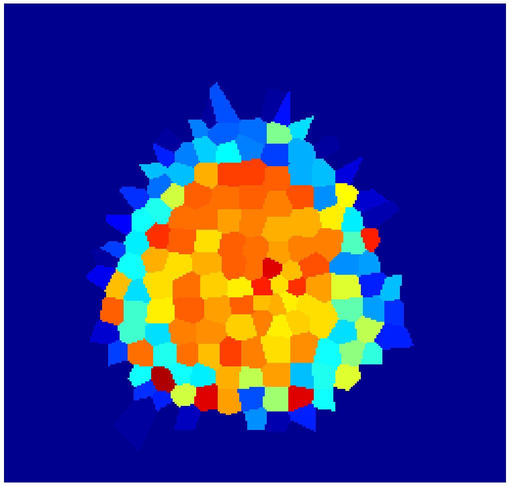
H1 with Estimation VQ



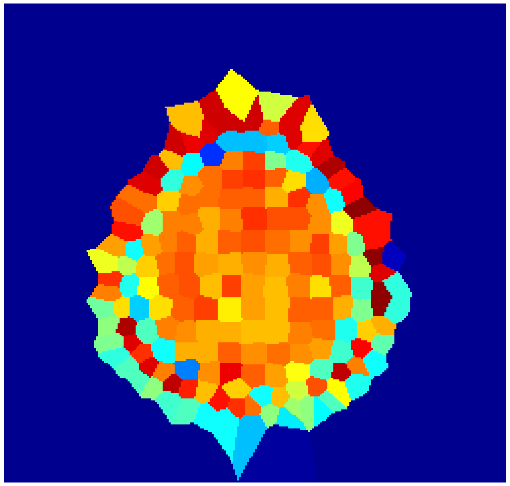
H0 with Detection VQ



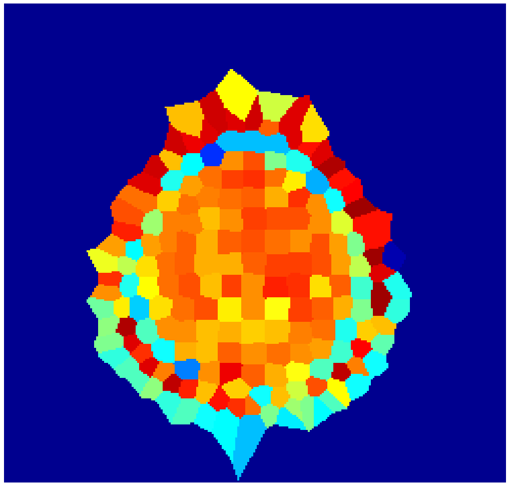
H1 with Detection VQ



H0 with Estimation VQ



H1 with Estimation VQ



Conclusions for Q/VQ for detection

- AUC criterion introduced: independent of detection threshold
- High rate Q/VQ analysis performed
- Good VQ's have cells aligned along contours of LR
- Optimal high rate Q/VQ strategies determined for various detection criteria
  1. One-sided discrimination exponent: Kullback Liebler divergence
  2. Two-sided discrimination exponent: α-divergence
  3. minimax exponent
  4. AUC exponent
- Application to longitudinal medical image databases is in progress