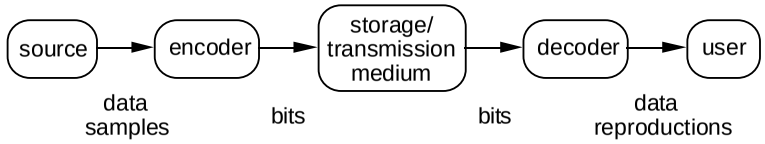
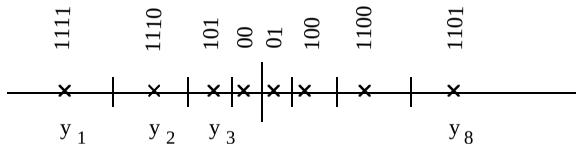
IV. SCALAR QUANTIZATION WITH VARIABLE-LENGTH CODING (SQ-VL)





- **Parameters of SQ-VL**
  - **size:** M (relatively unimportant; no reason for it to be power of 2; may be infinite.)
  - **thresholds** (for encoding)**:** $t_1,\ldots,t_{M-1}$, ($t_0 = x_{min}$, $t_M = x_{max}$; or $t_0 = -\infty$, $t_M = \infty$)
  - **levels** (for decoding)**:** $y_1,\ldots,y_M$
  - **binary codewords:** variable-length, <u>prefix</u> <u>code</u> $C_b = \{\underline{c}_1,\ldots,\underline{c}_M\}$, $l_i = l(\underline{c}_i)$
- Derivative characteristics:
  - **encoding rule:** $e(x) = \underline{c}_i$ when $t_{i-1} < x < t_i$
  - **decoding rule:** $d(\underline{b}) = y_i$ when $\underline{b} = \underline{c}_i$
  - **quantization rule:** $Q(x) = d(e(x)) = y_i$ when $t_{i-1} < x < t_i$

- Distortion:

$$D = E\,(X-Q(X))^2 = \int_{x_{min}}^{x_{max}} (x-Q(x))^2 f_X(x)\,dx = \sum_{i=1}^{M} \int_{t_{i-1}}^{t_i} (x-y_i)^2 f_X(x)\,dx$$

  Not affected by binary code
- Rate:

$$R = E\,l(e(X)) = \sum_{i=1}^{M} P_i\, l_i\,, \quad \text{where } P_i = \Pr(t_{i-1}<X<t_i) = \int_{t_{i-1}}^{t_i} f_X(x)\,dx$$

Notes:

- Higher probability cells should have shorter codewords.
- Rate is not affected by quantization levels.
- If Huffman or Shanon code is designed for the probabilities $P_1,\ldots,P_M$, then

$$H \le R < H + 1$$

  where $H = H(Q(X)) = H(P) = -\sum_{i=1}^{M} P_i \log_2 P_i$

  is the quantizer (output) entropy, i.e. the entropy of the $Y = Q(X)$ or $I = e(X)$.
- H depends only on the thresholds (not the levels).
- SQ-VL is often called *scalar quantization with entropy coding* (SQ-EC).
- H is often used as the definition of rate for SQ-VL, since it is an approximation to rate when the VLC is well designed.
- M is not a very important parameter because two quantizers can have very different M's, yet have approximately same R and D. For example, D and R will change imperceptibly if very very large thresholds and levels are added. M can even be infinity.

- Questions:
  - How to design/optimize SQ with VLC?
    (a) Iterative design algorithms?
    (b) High-resolution analysis for large rates?
  - What is the OPTA? The OPTA is defined below

$$\delta_{sq,vl}(R) \triangleq \text{least dist'n of SQ w VLC with rate } R \text{ or less}$$

  Can use either (a) or (b) to compute the OPTA.

- What should we expect?
  - $\delta_{sq,vl}(R) < \delta_{sq}(R)$. But how much less is the SQ-VL OPTA?
  - When minimizing distortion of M-level SQ subject to VL rate constraint, we might want to compromise distortion somewhat in order to make cells more unequal in probability so VLC rate can be less.
  - Levels have no affect on rate so we should choose them to be centroids. It will be the thresholds that we compromise.

- A common design approach -- Given target rate R, design a scalar quantizer with smallest possible MSE among those with output entropy $\leq$ R. Then design a Huffman code for the resulting cell probabilities. This works well if R is large (e.g. $R \geq 3$), but not if R is small (e.g. $R << 1$), because a Huffman code cannot have rate less than 1.
- We'll discuss high-resolution analysis first because aside from leading to formulas for the OPTA function, it will also lead to a simple design procedure.

HIGH RESOLUTION ANALYSIS OF SQ-VL

- We seek high resolution, approximate formulas for distortion D and rate R of a scalar with variable-length coding when rate is high and distortion is small.

  We assume most cells are small, neighboring cells have similar sizes, and levels are at the center of the cells.

- Distortion: Recall Bennett's integral, which provices an approximate formula for D:

$$D \cong \frac{1}{12M^2} \int_{-\infty}^{\infty} \frac{f_X(x)}{\lambda^2(x)}\, dx$$

  where $\lambda(x)$ = point density.

- However, since M has little significance when using variable-length coding, consider a point density $\Lambda(x)$ that is not normalized by M. It is defined to be a function having the following properties:

  (1) $\int_a^b \Lambda(x)\, dx \cong$ # levels in [a,b], (2) $\int_{-\infty}^{\infty} \Lambda(x)\, dx \cong$ total number of levels.

  (3) $\Lambda(x) \geq 0$, all x, (4) $\Lambda(x) \cong \frac{1}{t_i - t_{i-1}}$, $t_{i-1} \leq x \leq t_i$ (for most x; i.e. w. high prob.)

- Restatement of Bennett's integral:

  If most cells are small, neighboring cells have similar sizes, and levels are approximately in the center of the cells, then

$$D \cong \frac{1}{12} \int_{-\infty}^{\infty} \frac{f_X(x)}{\Lambda^2(x)}\, dx$$

HIGH-RESOLUTION FORMULA FOR RATE, I.E. ENTROPY

- Rate R: Recall that when the VLC is optimized, $H \le R < H + 1$. Therefore, when R is large, we can use H as a good approximation, since $R \cong H$.
- High-resolution, approximate formula for H, i.e. rate, when most cells are small and neighboring cells have similar sizes:

$$\mathbf{H} \cong h + \int_{-\infty}^{\infty} f_X(x) \log_2 \Lambda(x)\, dx$$

where h is *differential entropy*, defined as

$$h = h(X) = h(f_X) \triangleq -\int_{-\infty}^{\infty} f_X(x) \log_2 f_X(x)\, dx$$

- Derivation:

$$H = -\sum_{i=1}^{M} P_i \log P_i = -\sum_{i=1}^{M} \left(\int_{t_{i-1}}^{t_i} f_X(x)\, dx \right) \log_2 \left(\int_{t_{i-1}}^{t_i} f_X(x)\, dx \right)$$

$$\cong -\sum_{i=1}^{M} \left(f_X(w_i) \int_{t_{i-1}}^{t_i} dx \right) \log_2 \left(f_X(w_i) \int_{t_{i-1}}^{t_i} dx \right) = -\sum_{i=1}^{M} f_X(w_i)\,(t_i - t_{i-1}) \log_2 (f_X(w_i)\,(t_i - t_{i-1}))$$

$$\cong -\sum_{i=1}^{M} \left(f_X(w_i) \log_2 \frac{f_X(w_i)}{\Lambda(w_i)} \right) \Delta_i, \quad \text{using } \Lambda(w_i) \cong \frac{1}{\Delta_i},\ \Delta_i = t_i - t_{i-1}$$

$$\cong -\int_{-\infty}^{\infty} f_X(x) \log_2 \frac{f_X(x)}{\Lambda(x)}\, dx$$

- Important example: uniform scalar quantizer with small cells and overload region with negligible probability: $H \cong h - \log_2 \Delta$ because $\Lambda(x) = \frac{1}{\Delta}$ for most x

OPTA OF SQ-VL VIA HIGH-RESOLUTION ANALYSIS

- We now seek an approximate formula for the OPTA of SQ-VL that will be accurate when R is large. Using the approximate formulas for distortion and

$$\delta_{sq,vl}(R) \cong \frac{1}{12} \min_{\Lambda} \int_{-\infty}^{\infty} \frac{f_X(x)}{\Lambda^2(x)}\, dx$$

where the minimum is over all choices of function $\Lambda(x)$ such that

(1) $\Lambda(x) \ge 0$ all x

(2) $h + \int_{-\infty}^{\infty} f_X(x) \log_2 \Lambda(x)\, dx \le R$

To find such a formula will need the following

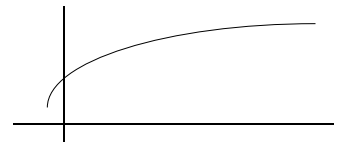
- Lower bound: If $h + \int_{-\infty}^{\infty} f_X(x) \log_2 \Lambda(x)\, dx \le R$, then

$$\log_2 \left(\int_{-\infty}^{\infty} \frac{f_X(x)}{\Lambda^2(x)}\, dx \right) \ge 2h - 2R$$

with equality if and only if $\Lambda(x) = 2^{R-h}$ wherever $f_X(x) > 0$.

To prove this we need the following:

- Definition: A function g is *concave* if for any z_1 and z_2 and any α, $0 < \alpha < 1$,



$$g(\alpha z_1+(1-\alpha)z_2) \geq \alpha g(\alpha z_1) + (1-\alpha) g(z_2)$$

  g is *strictly concave* if the inequality is always strict.

- Jensen's inequality:

  If Z is a random variable and g is a concave function, then

$$g(E[Z]) \geq E g(Z).$$

  If g is strictly concave, then

$$E g(Z) = g(E[Z])$$

  if and only if Z is constant with probability one.

  (If Z is a binary random variable, then Jensen's inequality is just a restatement of the definition of concave. Thus, Jensen's inequality can be viewed a direct extension of the definition of concave.)

- Proof of Lower Bound: Suppose $\Lambda(x)$ satisfies (1) and (2). Then

$$\log_2\left(\int_{-\infty}^{\infty} \frac{1}{\Lambda^2(x)} f_X(x)\, dx \right) = \log_2 E[Z], \qquad \text{where } Z = \frac{1}{\Lambda^2(X)}$$

$$\geq E[\log_2 Z], \qquad \text{using Jensen's ineq'y and fact that } \log_2 z \text{ is concave}$$

$$= \int_{-\infty}^{\infty} \log_2 \frac{1}{\Lambda^2(x)} f_X(x)\, dx = -2 \int_{-\infty}^{\infty} f_X(x) \log \Lambda(x)\, dx$$

$$\geq 2h - 2R \qquad \text{since} \qquad h + \int_{-\infty}^{\infty} f_X(x) \log_2 \Lambda(x)\, dx \leq R$$

  This demonstrates the lower bound.

  It is easy to check that equality holds if $\Lambda(x) = 2^{R-h}$ wherever $f_X(x) > 0$.

  Conversely, if equality holds, then equality must hold in the application of Jensen's inequality to $\log_2 Z$. Since $\log_2$ is strictly concave, this means $Z = 1/\Lambda^2(X)$ is constant with probability one, i.e. $\Lambda(x)$ is the same for all values of x such that $f_X(x) > 0$. Examining the second inequality above, we see that the value of
  $\Lambda(x)$ must be be 2^{R-h}.

- This completes the proof that $\log_2\left(\int_{-\infty}^{\infty} \frac{f_X(x)}{\Lambda^2(x)}\, dx \right) \geq 2h - 2R$ and that equality is possible. Therefore,

$$\int_{-\infty}^{\infty} \frac{f_X(x)}{\Lambda^2(x)}\, dx \geq 2^{2h}\, 2^{-2R}$$

- Moreover we showed that the left hand side can be as small as the right side. In particular this will happen if $\Lambda(x)$ is a constant, which happens if the quantizer is uniform with an infinite number of levels.

  It is also sufficient that $\Lambda(x)$ is constant on the set of x where $f_X(x) > 0$ and essentially sufficient if $\Lambda(x)$ is constant on a set of high probability. Thus, a USQ with a large but finite numer of cells covering all but a region with negligible probability will suffice.

  It follows from the lower bound and its achievability that

- For large R, the OPTA of SQ-VL is

$$\delta_{sq,vl}(R) \cong \frac{1}{12}\sigma^2 \eta\, 2^{-2R}$$

and

$$S_{sq,vl}(R) \cong 6.02\,R - 6.02\,h + 10\log_{10} 12\,\sigma^2$$

$$= 6.02R + 10\log_{10}\frac{12}{\eta}$$

where

$$\eta = \frac{1}{\sigma^2} 2^{2h},$$

$$h = -\int_{-\infty}^{\infty} f_X(x)\log_2 f_X(x)\,dx$$

or equivalently $\eta = 2^{2\tilde{h}}$, where $\tilde{h}$ = differential entr. of X/σ

Notes:

- Again, the SNR OPTA increases at 6 dB per bit.
- For large rate R, the optimal point density is constant, $\Lambda(x) = 1/\Delta$, and a uniform quantizer with infinitely many levels and step size Δ is optimal, i.e

$$D_{usq} \cong \frac{\Delta^2}{12} \cong \delta_{sq,vl}(R),$$

  where the step size Δ is chosen so that $\frac{\Delta^2}{12} = \frac{1}{12} 2^{2h}\, 2^{-2R}$.

- For all sufficiently large values of M, the best performance of SQ-VL with M levels is as good as an infinite number of levels, i.e.

$$\delta_{usq,vl,M}(R) \cong \delta_{sq,vl}(R)\ .$$

  So we can use USQ with a finite number of levels.

- How large must M be? Experience suggests

$$M \cong 2 \times 2^R$$

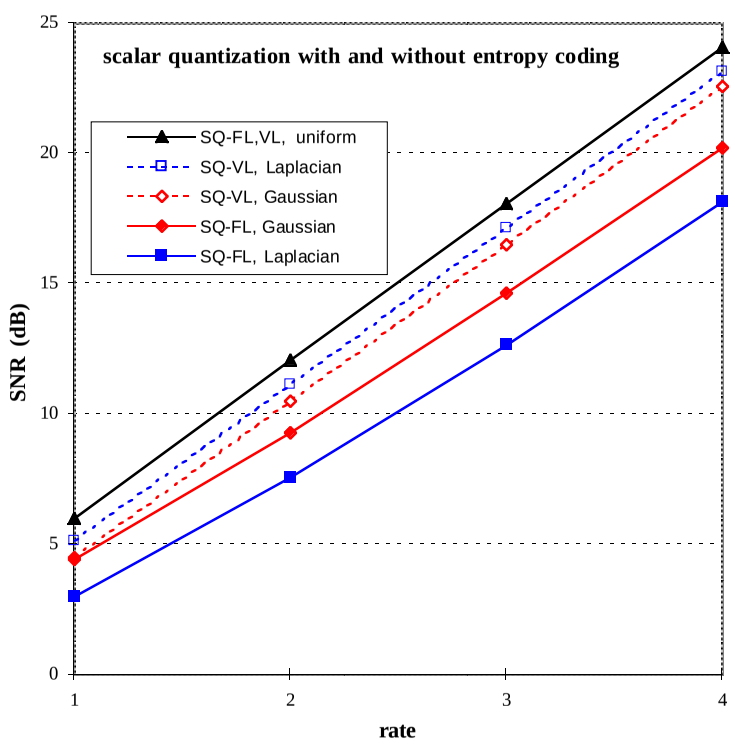
  is the smallest value that works. Larger values are also fine. For example, see Goblick and Holsinger, Fig. 2, IEEE Inform. Thy, April 1967, which shows entropy vs SNR for USQ's with M levels as Δ varies, for M = 10, 20, 40. More discussion later.

- Nonuniform SQ-VL can do somewhat better than USQ-VL at low rates. There are algorithms for optimizing nonuniform SQ-VL at low rates.

Comparison: USQ-VL vs. USQ-FL

- USQ-VL is much better than USQ-FL. Its SNR increases as 6R + constant, whereas SNR of USQ-FL increases as $6R - \log_{10}R$ + constant.
- Why so much better? Since the rate of USQ-VL need not increase with the number of levels, we can choose M so large that overload distortion is negligible. Therefore, we don't have to compromise granular distortion to reduce overload distortion, as we do with USQ-FL.

COMPARISON: SQ-VL VS. SQ-FL



COMPARISON: SQ-VL VS. SQ-FL

SQ-FL

$$\delta_{sq,fl}(R) \cong \frac{1}{12}\ \sigma^2 \beta 2^{-2R}$$

$$S_{sq,fl}(R) \cong 6.02\,R + 10\log_{10}\frac{12}{\beta}$$

$$\Lambda^*(x) = \text{constant}$$

SQ-VL

$$\delta_{sq,vl}(R) \cong \frac{1}{12}\sigma^2 \eta 2^{-2R}$$

$$S_{sq,vl}(R) \cong 6.02R + 10\log_{10}\frac{12}{\eta}$$

$$\lambda^*(x) = c\, f_X^{1/3}(x)$$

	uniform	**Gaussian**	**Laplacian**
β	12	$6\sqrt{3}\pi$=32.65	54
$10\log_{10}\frac{12}{\beta}$	0 dB	-4.347 db	-6.532 dB
$10\log_{10}\frac{12}{\eta}$	0 dB	-1.533 dB	-0.904 dB
VL gain	0 dB	2.81 dB	5.63 dB
$\eta = 2^{2h}/\sigma^2$	12	$2\pi e$=17.08	$2e^2$=14.78
h	$\frac{1}{2}\log_2 12\sigma^2$	$\frac{1}{2}\log_2 2\pi e\sigma^2$	$\frac{1}{2}\log_2 2e^2\sigma^2$

- As expected, optimal SQ-VL is at least as good as optimal SQ-FL. Indeed it is better, except for uniform densities.
- Fact: $\eta \le \beta$. Equality holds iff $f_X(x)$ is uniform. Proof comes later.
- Also as expected, opt'l pt. dens. for SQ-VL makes cells have more disparate probabilities than opt'l pt dens. for SQ-FL. (Pr(cell cont'g x) is proportional to $f_X(x)$ for SQ-VL, whereas it is proportional to $f_X^{2/3}(x)$ for SQ-FL.)
- Why is SQ-VL so much better than SQ-FL for Laplacian?

  The Laplacian pdf has a heavy tail and its SQ-FL OPTA is heavily influenced by the need to spread the levels out far enough to reduce what would otherwise be a large overload distortion. Thus with SQ-FL one must signficantly compromise granular distortion to reduce overload distortion. On the other hand, with SQ-VL, we don't have to about the overload distortion because we can get good performance with $M = \infty$ and no overload region whatsoever.

PROPERTIES OF DIFFERENTIAL ENTROPY:

(1) $\eta \le \beta$ Equality holds iff $f_X(x)$ has the same value wherever it is not zero, e.g. if it is uniform.

Proof:

$$h = -\int f_X(x) \log_2 f_X(x)\, dx = \frac{3}{2}\int f_X(x) \log_2 f_X^{-2/3}(x)\, dx$$

$$= \frac{3}{2} E \log_2 Y, \quad \text{where } Y = f_X^{-2/3}(X)$$

$$\le \frac{3}{2} \log_2 EY \quad \text{by Jensen's ineq. } (\log_2 \text{ is concave})$$

$$= \frac{3}{2} \log_2 \int f_X(x) f_X^{-2/3}(x)\, dx = \frac{1}{2} \log_2 \left(\int f_X^{-1/3}(x)\, dx \right)^3 = \frac{1}{2} \log_2 \sigma^2 \beta$$

$$\Rightarrow \sigma^2 \eta = 2^{2h} \le \sigma^2 \beta$$

Since $\log_2$ is strictly concave, equality holds iff Y is constant w.p.1; i.e. iff $P(f_X^{-2/3}(X) = c) = 1$ for some some c, i.e. iff $f_X(x)$ has the same value wherever it is not zero.

(2) Uniform density: if f_X uniform on [a,b] ,

$$h = \log_2 (b-a) = \frac{1}{2} \log_2 12\sigma^2 \quad \text{and} \quad \eta = 12$$

(3) Laplacian density: $h = \frac{1}{2} \log_2 2e^2\sigma^2$ and $\eta = 2e^2 = 14.8$

(4) Gaussian density: $h = \frac{1}{2} \log_2 2\pi e\, \sigma^2$ and $\eta = 2\pi e = 17.08$

(5) For any density with variance σ^2:

$$h \le \frac{1}{2} \log_2 2\pi e \sigma^2,$$

and so $\eta \le 2\pi e$ for any density.

Equality iff X is Gaussian.

$\Rightarrow$ Gaussian is hardest density to quantize with SQ-VL.

<u>Proof:</u> Let $f_X(x)$ be an arbitrary density with variance σ^2 and let $g(x)$ be the Gaussian density with same mean and variance.

We make the proof in two steps.

(a) $h(f) \le -\int_{-\infty}^{\infty} f(x) \log_2 g(x)\, dx$:

$$-\int_{-\infty}^{\infty} f(x) \log_2 g(x)\, dx - h(f) = -\int_{-\infty}^{\infty} f(x) \log_2 \frac{g(x)}{f(x)}\, dx$$

$$\ge -\int_{-\infty}^{\infty} f(x) \left(\frac{g(x)}{f(x)} - 1 \right) dx \frac{1}{\ln 2}, \quad \left(\log_2 z \le (z-1) \frac{1}{\ln 2} \right)$$

$$= -\int_{-\infty}^{\infty} g(x)\, dx \frac{1}{\ln 2} + \int_{-\infty}^{\infty} f(x)\, dx \frac{1}{\ln 2} = -1 + 1 = 0$$

(b) $-\int_{-\infty}^{\infty} f(x) \log_2 g(x)\, dx = \frac{1}{2} \log_2 2\pi e \sigma^2$:

$$-\int_{-\infty}^{\infty} f(x)\ \log_2 g(x)\ dx \ = \ -\int_{-\infty}^{\infty} f(x)\ \log_2 \left(\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\}\right) dx$$

$$= -\int_{-\infty}^{\infty} f(x)\ \log_2 \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) dx - \int_{-\infty}^{\infty} f(x) \left(-\frac{(x-m)^2}{2\sigma^2}\right) \log_2 e\ dx$$

$$= \log_2 \sqrt{2\pi\sigma^2} + \frac{1}{2} \log_2 e \ = \ \frac{1}{2} \log_2 2\pi e\sigma^2$$

Together (a) and (b) imply $h(f) \le \frac{1}{2} \log_2 2\pi e\sigma^2$

Equality holds iff it holds in (a), which happens iff $\frac{g(x)}{f(x)} = 1$, i.e. iff f(x) is Gaussian.

(6) If $Y = aX+b$, then $h_Y = h_X + \log_2 |a|$ and $\eta_Y = \eta_X$.

Unlike entropy H, differential entropy h is affected by scaling. It is not affected by shifting. Like the Panter-Dite factor β, η is not affected by scaling or shifting.

(7) h can be positive or negative.

For example, let $|a|$ be very small in the previous fact.
h negative simply means that $\eta < 1$.

(8) h (and η) can be infinite, but usually it is finite.

In this case $\delta_{sq,vl}(R)$ does not decrease as 2^{-2R}.

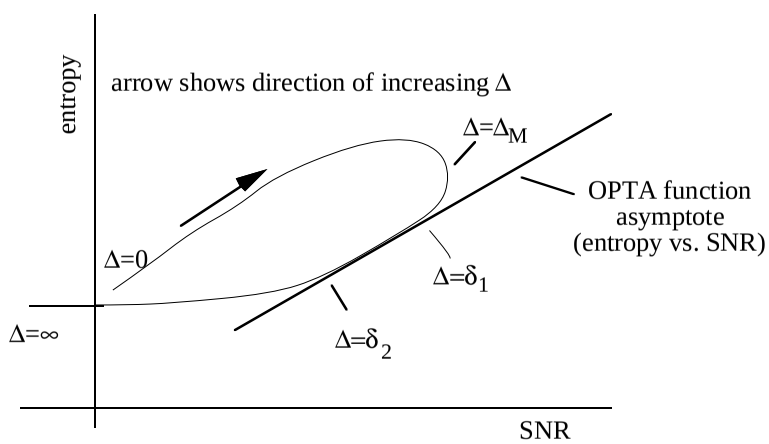
ENTROPY VS SNR OF USQ WITH M LEVELS AS Δ RANGES FROM 0 TO ∞.

- Case 1, Theshold at the origin:

  For very large Δ, SNR << 0 ($D_{overload} \cong 0$, but D_{gran} is huge) and H =1 (two innermost levels have probability 1/2).

  As Δ decreases to δ_1, entropy increases, D_{gran} decreases and the $D_{overload}$ remains negligible.

  As Δ decreases from δ_1 to Δ_M (= MSE minimizing Δ), overall dist'n continues to decrease but $D_{overload}$ becomes significant.



As Δ decreases from Δ_M, the overall dist'n increases and for a little while entropy increases, but then it decreases back towards 1 as $\Delta \to 0$ (two outermost levels have probability 1/2). Between δ_1 and δ_2 the performance lies approximately on the opta function for SQ-VL.

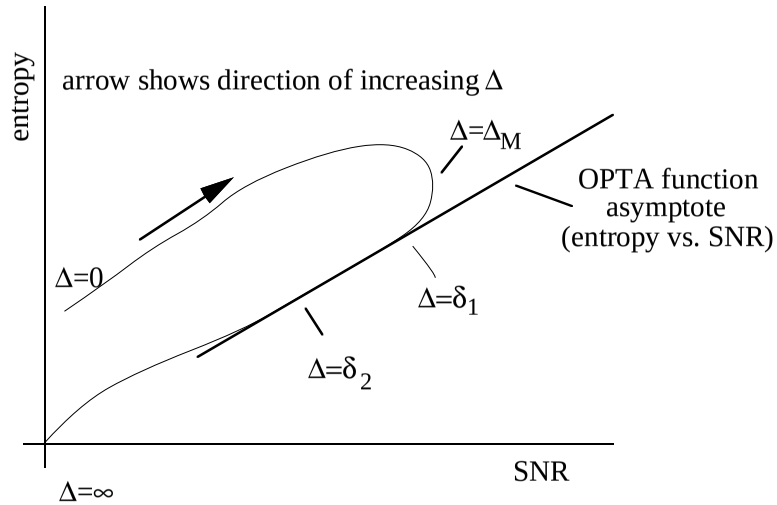
ENTROPY VS SNR OF USQ WITH M LEVELS AS Δ RANGES FROM 0 TO ∞

- Case 2, Level at the origin

  For very large Δ, SNR $\cong 0$ ($D_{overload} \cong 0$, but $D_{gran} = \sigma^2$) and H = 0 (level at 0 has probability 1).

  As Δ decreases to δ_1, the entropy increases, D_{gran} decreases and the $D_{overload}$ remains negligible.

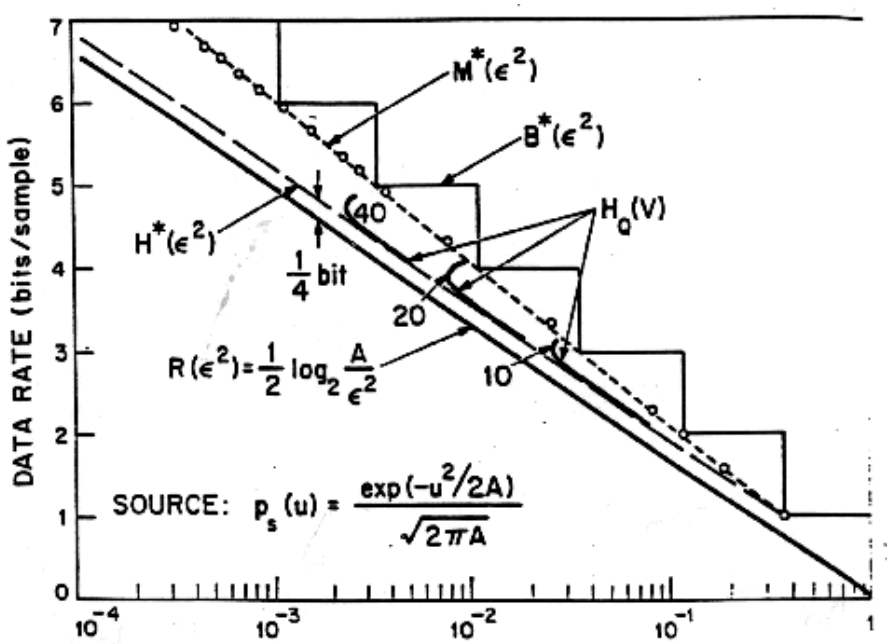
  As Δ decreases from δ_1 to Δ_M (= MSE minimizing Δ), overall dist'n continues to decrease but $D_{overload}$ becomes significant.



  As Δ decreases from Δ_M, overall dist'n increases and for a little while entropy increases, but then it decreases back towards 1 as $\Delta \to 0$ (two outermost levels have probability 1/2). Between δ_1 and δ_2 the performance lies approximately on the opta function for SQ-VL.

- Note: If small rate and small M desired, e.g. R ≤ 2, need level at the origin, so use M odd, so quantizer can be symmetric.

PLOT FROM GOBLICK & HOLSINGER, IEEE TRANS. INFORM. THY., APRIL 1967



Rate vs. D/σ^2

$H_Q(V)$ is output entropy of USQ with M = 10, 20, 40 levels and step size that varies.

$B^*(\varepsilon^2)$ is the rate of USQ with FLC ($\lceil \log_2 M \rceil$)