577

Upper Bounds to the Asymptotic Performance of Block Quantizers

JAMES A. BUCKLEW, MEMBER, IEEE

Abstract—Upper bounds to the asymptotic performance of block quantizers with difference distortion measures are derived. In many cases, these upper bounds approach known lower bounds as the block length of the quantizer approaches infinity. A condition for the optimal point density function of the output levels is derived. It is shown to particularize to a known result of Gersho. The behavior of the bounds for large block lengths is investigated.

I. Introduction

ZADOR [1] and Gersho [2] consider the problem of minimizing the average distortion for block or vector quantizers. They derive an expression for the optimal asymptotic performance of a vector quantizer when the distortion measure is the rth power of the Euclidean distance. This expression includes a parameter which depends only on r and k (the dimension of the vector to be quantized), but which has an unknown form in all but a few special cases. Zador, making use of random coding arguments, derives upper bounds to this parameter. In the same manuscript he derives a sphere packing lower bound. An interesting fact about these results is that the upper bound asymptotically approaches the lower bound as $k \to \infty$.

In a more recent paper, Yamada *et al.* [3] calculate lower bounds to the asymptotic (high rate) performance of block quantizers with more general difference distortion measures. They then proceed to show how these lower bounds are, in fact, tighter in many cases than the Shannon lower bound to the rate distortion function.

In this paper our objectives are twofold. First, we develop upper bounds for the same distortion measures that are treated in [3]. In many cases we can show that, as the block length k approaches infinity, these upper bounds approach the lower bounds. Second, we investigate the limiting case of $k \to \infty$ and discuss a possible application to the implementation of block quantizers. We devote the final section to a few observations and remarks.

II. Upper Bound Derivation

We employ the same notation and definitions found in [3]. We reproduce the relevant notation below for convenience. Given $G \subset \mathcal{R}^k$, define $V(G) = \int_G dx$. Let $V_k = V(\{u \in \mathcal{R}^k: \|u\| \le 1\})$, where $\| \cdot \|$ is an arbitrary semi-

norm on $\mathcal{R}^k$. Obviously, V_k is the volume of a sphere of radius 1. The volume of a sphere of radius a is given by $V_k a^k$. The distortion resulting from reproducing a vector x as y is defined as a difference distortion measure by $L(x - y)$, where

a) $L(0) = 0$,

b) $L(u) \le L(x)$ if and only if $\|u\| \le \|x\|$,

and

c)
$$M_k(v) = \frac{1}{V_k} \int_{\|u\| \le 1} L\left(V_k^{-1/k} v u\right) du$$
$$= \frac{1}{v^{+k}} \int_{\|x\| \le v V_k^{-1/k}} L(x) \, dx$$

is assumed to be convex $\cup$ in V.

Consider the asymptotic distortion produced by a random quantizer. A random quantizer is a quantizer wherein the N output levels are chosen to be samples from some k-dimensional probability density function $\lambda(x)$. The k-dimensional data vectors x are assumed to be samples from another k-dimensional probability density $p(x)$. Both $p(x)$ and $\lambda(x)$ are assumed to be sufficiently smooth so that they may be considered to be constant over small bounded sets in $\mathcal{R}^k$. For large enough N we will assume that $p(x)$ has support only on a compact set K.

Consider a data point $x \in \mathcal{R}^k$ to be quantized. We wish to pick the output level y that minimizes the distortion measure $L(x - y)$. From property b) this means choosing the value y that minimizes the seminorm $\|x - y\|$. Suppose that y exists on a sphere of radius r. Since $p(x)$ is smooth, if N is very large we may suppose that y is uniformly distributed on the sphere. We wish to calculate the expected value of the distortion if the closest y is distributed uniformly on a sphere of radius r. Since $\int_{\|x\| \le r} L(x) \, dx$ is a convex $\cup$ function of r, the expected distortion must be bound above by the limit as $\epsilon \downarrow 0$ of

$$I_{r,\epsilon} = \frac{\int_{\|x\| \le r+\epsilon} L(x) \, dx - \int_{\|x\| \le r} L(x) \, dx}{V_k\left[(r + \epsilon)^k - r^k\right]},$$

but $(r + \epsilon)^k - r^k \simeq \epsilon k r^{k-1}$ for $\epsilon \ll r$. Also

$$M_k\left((r + \epsilon)V_k^{1/k}\right) = (r + \epsilon)^{-k} V_k^{-1} \int_{\|x\| \le r+\epsilon} L(x) \, dx.$$

Hence

$$I_{r,\epsilon} \cong \frac{M_k\big((r+\epsilon)V_k^{1/k}\big)V_k(r+\epsilon)^k - M_k\big(rV_k^{1/k}\big)V_k r^k}{V_k \epsilon k r^{k-1}}$$

$$= \frac{r}{k}\frac{M_k\big((r+\epsilon)V_k^{1/k}\big) - M_k\big(rV_k^{1/k}\big)}{\epsilon}$$

$$+ M_k\big((r+\epsilon)V_k^{1/k}\big).$$

Take the limit as $\epsilon \to 0^+$ to obtain

$$\lim_{\epsilon \to 0^+} I_{r,\epsilon} = I_r = \frac{r}{k}V_k^{1/k}M_k'^+\big(rV_k^{1/k}\big) + M_k\big(rV_k^{1/k}\big)$$

where $M_k'^+(\cdot)$ indicates the right derivative of $M_k(\cdot)$ and always exists for convex functions [4].

We need to calculate the probability density function of r, the distance of the closest output level to x. The probability that one output level will fall in a sphere of radius r about the point x is approximately (for large N) $r^k V_k \lambda(x)$. Therefore, the probability density function of r is (using order statistics) [5] $\int_{\Re^k} N[1 - r^k V_k \lambda(x)]^{N-1} k r^{k-1}\lambda(x) \cdot V_k p(x)\, dx$. Hence an upper bound to the average distortion produced by a random quantizer must be

$$\int\int I_r N\big[1 - r^k V_k \lambda(x)\big]^{N-1} k r^{k-1} V_k \lambda(x)p(x)\, dx\, dr.$$

This must be an upper bound to the distortion rate function (for large rates) because the output levels were chosen randomly. Since $M_k(v)$ is convex $\cup$ we know that $M_k(v) \geq M_k(v_0) + M'^+(v_0)(v - v_0)$. Therefore,

$$I_r = M_k\big(rV_k^{1/k}\big) + \frac{rV_k^{1/k}}{k}M_k'^+\big(rV_k^{1/k}\big)$$

$$\leq M_k\Big(rV_k^{1/k}\Big(1 + \frac{1}{k}\Big)\Big).$$

Hence the distortion is bound above by

$$\int\int M_k\Big(rV_k^{1/k}\Big(1 + \frac{1}{k}\Big)\Big)N\big[1 - r^k V_k \lambda(x)\big]^{N-1}$$

$$\cdot k r^{k-1} V_k \lambda(x)p(x)\, dx\, dr.$$

Make a change of variable $u = r^k V_k \lambda(x)$, $du = k r^{k-1}V_k \lambda(x)\, dr$, $u/\lambda(x) = r^k V_k$, and note that $u \leq 1$ allows us to rewrite our upper bound as

$$\int_{\Re^k}\int_0^1 M_k\left(\left(\frac{u}{\lambda(x)}\right)^{1/k}\frac{k+1}{k}\right)$$

$$\cdot N[1 - u]^{N-1}\, du\, p(x)\, dx = D_u.$$

Now note that $N[1 - u]^{N-1}$ is the density function of the smallest of N samples taken from a uniform density on the interval $(0, 1)$. Hence $E\{u\} = 1/(N + 1)$ var $\{u\} \leq 1/(N + 1)^2$. The lower bound derived in [3] is $D_L = \int_{\Re^k} M_k((N\lambda(x))^{-1/k})p(x)\, dx$. The closeness of D_u to D_L depends entirely on the properties of $M_k(\cdot)$.

Let us assume that $M_k(\cdot)$ is p-varying near zero for some $p \geq 0$; that is, for all $t > 0$, $\lim_{x \downarrow 0}(M(tx)/M(x)) = t^p$. This is not a very restrictive assumption. If for all t, $\lim_{x \downarrow 0}(M(tx)/M(x)) = h(t)$ for some function h, then it

is necessarily true that $h(t) = t^p$ for some $-\infty \leq p \leq \infty$. Suppose further that $\lambda(x) \geq \beta > 0$. Let $c > 0$ be an arbitrary constant. Consider the following quantity:

$$\int_0^1 \frac{M_k\big(cu^{1/k}\big)}{M_k\big(c/n^{1/k}\big)}n(1 - u)^{n-1}\, du$$

$$= \int_0^n \frac{M_k\big(cV^{1/k}/n^{1/k}\big)}{M_k\big(c/n^{1/k}\big)}\Big(1 - \frac{V}{n}\Big)^{n-1}\, dv$$

$$= \int_0^A(\cdot) + \int_A^n(\cdot).$$

We wish to show that for all $\epsilon > 0$, one may find an $A > 1$ large enough such that

$$\Gamma(p/k + 1) > \int_0^A(\cdot) > \Gamma(p/k + 1) - \epsilon \qquad (1)$$

for all large enough n, and

$$\lim_{n \to \infty}\sup\int_A^n(\cdot) < \epsilon. \qquad (2)$$

For $V \in (0, A)$ we have

$$\frac{M_k\big(cV^{1/k}/n^{1/k}\big)}{M_k\big(c/n^{1/k}\big)} \leq \frac{M_k\big(cA^{1/k}/n^{1/k}\big)}{M_k\big(c/n^{1/k}\big)}$$

which is uniformly bounded for all $n \geq 1$. Since $(1 - (V/n))^{n-1} \leq (1 - V/n)^n/(1 - (A/n)) \leq e^{-v}/(1 - (A/n))$ is uniformly bounded for all $n > A$, we may apply the dominated convergence theorem to obtain

$$\lim_{n \to \infty}\int_0^A(\cdot)$$

$$= \int_0^A \lim_{n \to \infty}\frac{M_k\big(cV^{1/k}/n^{1/k}\big)}{M_k\big(c/n^{1/k}\big)}\Big(1 - \frac{V}{n}\Big)^{n-1}\, dv$$

$$= \int_0^A V^{p/k}e^{-v}\, dv < \int_0^\infty V^{p/k}e^{-v}\, dv = \Gamma\Big(\frac{p}{k} + 1\Big).$$

For A sufficiently large this limit is larger than $\Gamma(p/k + 1) - \epsilon$. This proves (1).

For all $1 < t \leq 2$ and all $x < \delta$ we have $M_k(tx)/M_k(x) < t^{p+1}$. Thus if $A < V < n(2\delta/c)^k$ we find

$$\frac{M_k\big(CV^{1/k}/n^{1/k}\big)}{M_k\big(C/n^{1/k}\big)}$$

$$\leq \frac{M_k\big(CV^{1/k}/n^{1/k}\big)}{M_k\big(CV^{1/k}/2n^{1/k}\big)}$$

$$\cdot \frac{M_k\big(CV^{1/k}/2n^{1/k}\big)}{M_k\big(CV^{1/k}/4n^{1/k}\big)} \cdots \frac{M_k\big(CV^{1/k}/2^{m-1}n^{1/k}\big)}{M_k\big(CV^{1/k}/2^m n^{1/k}\big)}$$

where

$$2^{m-1} < V^{1/k} \leq 2^m \leq (2^{p+1})^m < (2^{p+1})((1/k)\log V + 1)$$

$$= 2^{p+1}V^{(p+1)/k}.$$

Assume without loss of generality that $2\delta/c < 1$. Then

$$\int_A^n (\cdot) = \int_A^{n(2\delta/c)^k} (\cdot) + \int_{n(2\delta/c)^k}^n (\cdot)$$

$$\le 2^{P+1} \int_A^\infty V^{(p+1)/k} e^{-v(n-1)/n} \, dv$$

$$+ \int_{n(2\delta/c)^k}^n \frac{M_k(c) e^{-v(n-1)/n}}{M_k(c/n^{1/k})} \, dv.$$

The first term is smaller than $\epsilon/2$ by choice of A (if $n \ge 2$). It is known that p-varying functions satisfy $M_k(x) = x^{p(1+0(1))}$ as $x \downarrow 0$. Thus

$$M_k(c/n^{1/k}) = \left(\frac{c}{n^{1/k}} \right)^{p(1+0(n^{-1/k}))},$$

so that the last integral is not larger than

$$M_k(c) \left(\frac{n^{1/k}}{c} \right)^{p(1+O(n^{-1/k}))} \int_{n(2\delta/c)^k}^\infty e^{-v/2} \, dv$$

$$< (\text{constant}) n^{2p/k} e^{-(2\delta/c)^k n/2}, \qquad n \ge 2,$$

for all n large enough; clearly, this bound goes to zero as $n \to \infty$. This proves (2). We now have that

$$\lim_{N \to \infty} \int \frac{M_k(cu^{1/k})}{M_k(c/N^{1/k})} N(1-u)^{N-1} \, du = \Gamma\left(\frac{p}{k} + 1 \right).$$

If $\lambda(x) > B > 0$ the above limit is uniform in x. Let $\epsilon > 0$. We then have $\exists N_0 \ni : N > N_0 \Rightarrow$

$$1 - \epsilon \le \frac{1}{\Gamma\left(\frac{p}{k} + 1 \right)} E_u \left\{ \frac{M_k\left(\frac{(k+1)u^{1/k}}{k\lambda(x)^{1/k}} \right)}{M_k\left(\frac{k+1}{k\lambda(x)^{1/k} n^{1/k}} \right)} \Big| x \right\}$$

$$\le 1 + \epsilon$$

$$\Gamma\left(\frac{p}{k} + 1 \right)(1 - \epsilon) M_k\left(\frac{k+1}{k\lambda(x)^{1/k} n^{1/k}} \right)$$

$$\le E_u \left\{ M_k\left(\frac{k+1u^{1/k}}{k\lambda(x)^{1/k}} \right) \Big| x \right\}$$

$$\le \Gamma\left(\frac{p}{k} + 1 \right)(1 + \epsilon) M_k\left(\frac{k+1}{k\lambda(x)^{1/k} n^{1/k}} \right).$$

Taking the expectation in x, dividing through and taking the limit as $N \to \infty$ we find

$$\lim_{N \to \infty} \frac{D_u}{\Gamma\left(\frac{p}{k} + 1 \right) \int M_k\left(\frac{k+1}{k\lambda(x)^{1/k} N^{1/k}} \right) p(x) \, dx} = 1.$$

Hence we may use the denominator of the above expression as an upper bound for the expected distortion in the high rate regions. We denote this upper bound as D_u' or

$$D_u' \triangleq \Gamma\left(\frac{p}{k} + 1 \right) \int M_k\left(\frac{k+1}{k} \left(\frac{1}{N\lambda(x)} \right)^{1/k} \right) p(x) \, dx.$$

Making use of this last equation, let us find the $\lambda(x)$ that minimizes D_u'. We desire to minimize

$$\int M_k\left(\frac{c}{\lambda(x)^{1/k}} \right) p(x) \, dx,$$

using a variational approach. Let $\lambda(x) = \lambda_0(x) + \epsilon f(x)$, where $\lambda_0(x)$ is the optimum point density and $\epsilon f(x)$ is an arbitrary variation. Obviously, $\int f(x) \, dx = 0$. The following condition must be satisfied

$$\frac{\partial}{\partial \epsilon} \int M_k\left(\frac{c}{(\lambda_0(x) + \epsilon f(x))^{1/k}} \right) p(x) \, dx \Big|_{\epsilon=0} = 0.$$

Assuming we can pull the derivative through the integral implies

$$\int M_k'\left(\frac{c}{\lambda_0(x)^{1/k}} \right) \frac{f(x)p(x)}{\lambda_0(x)^{1+1/k}} \, dx = 0.$$

The only solution that will work for arbitrary $f(x)$ is if the following is a constant ($p(x) - $ a.e.):

$$M_k'\left(\frac{c}{\lambda_0(x)^{1/k}} \right) \frac{p(x)}{\lambda_0(x)^{1+1/k}}.$$

As an example, for the distortion measure given by $L(x - y) = \|x - y\|_2^r (r \ge 1)$, we have from [3] that

$$M_k(v) = \frac{kv^r}{k+r} V_k^{-r/k}.$$

The condition implies then that $\lambda_0(x)$ is proportional to $p(x)^{k/(k+r)}$, a fact derived in [1], [2] using Holder's inequality. One should note that for large k $\lambda_0(x) \cong p(x)$. We will have more to say about this in the next section.

III. LARGE DIMENSIONAL BEHAVIOR

We will need an additional assumption of the sequence of functions $M_k(')$. We assume that $M_k(v)/k$ converges pointwise to some other function $M(v)$. Single letter distortion measures satisfy this constraint. We note that for distortion measures of the form $L(x - y) = \|x - y\|_v^r$, only when $r = v$ do we have a single letter distortion measure. The $1/k$ factor is added so that we may consider the distortion per dimension (the relevant quantity for distortion-rate theory comparisons). Since the $M_k(v)$ are convex, this implies $M(v)$ must be convex. It can also be shown that $M_k(v)/k$ converges uniformly to $M(v)$ on compact sets [4]. If we may neglect $p(x)$ outside a compact set K then the Lipschitz condition for convex functions implies that $(1/k)(D_u' - D_L) \to 0$. Suppose $p(x)$ is a restriction of a stationary ergodic measure. Now $p(x)$ induces a measure μ on $\mathcal{R}^k$: $\mu(A) = \int_A p(x) \, dx$ for every Borel measurable set A in $\mathcal{R}^k$. If the measure induced by $\lambda(x)$ is absolutely continuous with respect to μ then subject to some mild regularity conditions [6], we have $-k^{-1} \ln \lambda(x) \xrightarrow{\mu} \lim_{k \to \infty} k^{-1} \int p(x) \ln \lambda(x) \, dx$. Since e^x is continuous, $\exp(-k^{-1} \ln \lambda(x)) \xrightarrow{\mu} \exp(\lim_{k \to \infty} k^{-1} \int p(x) \ln \lambda(x) \, dx)$ [7, p. 31]. Finally, since $M_k(v)/k \to M(v)$

uniformly on compact sets, it is continuous, and hence analogous to [7, p. 31] we have

$$\frac{1}{k} M_k\left(\frac{e^{-(1/k)\ln\lambda(x)}}{2^b}\right) \xrightarrow{\mu} M\left(\frac{e^{-\lim_{k\to\infty}(1/k)\int p(x)\ln\lambda(x)}}{2^b}\right)$$

where $b \triangleq$ the rate in bits per sample used to quantize the data ($N = 2^{bk}$). We have that

$$\lim_{k\to\infty} E_\mu\left\{\frac{1}{k} M_k\left(\frac{e^{-(1/k)\ln\lambda(x)}}{2^b}\right)\right\}$$

$$= M\left(\frac{e^{-\lim_{k\to\infty}(1/k)\int p(x)\ln\lambda(x)\,dx}}{2^b}\right)$$

if and only if the random variables

$$\frac{1}{k} M_k\left(\frac{e^{-(1/k)\ln\lambda(x)}}{2^b}\right)$$

are uniformly integrable.

A sufficient condition [7] for uniform integrability is that for some $\epsilon \to 0$

$$\sup_k \frac{1}{k} M_k\left\{\left(\frac{e^{-(1/k)\ln(x)}}{2^b}\right)\right\}^{1+\epsilon} < \infty.$$

Gersho [2] shows that this condition is met for $L(x) = \|x\|_2^r(\|\cdot\|_2$ is the l_2 or Euclidean distance norm) if $\|p(x)\|1/(1+r_0) < \infty$ ($r_0 > r \geq 1$). Suppose we do have uniform convergence of these random variables. Then the distortion for very large k may be approximated by

$$D = M\left(\frac{1}{2^b} e^{-\lim_k(1/k)\int p(x)\ln\lambda(x)\,dx}\right).$$

$M(\cdot)$ is nondecreasing, which implies that in order to minimize the large dimensional asymptotic distortion, it is sufficient to minimize the argument of the $M(\cdot)$ function. However, since

$$-\frac{1}{k}\int p(x)\ln\lambda(x)\,dx \geq -\frac{1}{k}\int p(x)\ln p(x)\,dx;$$

this implies that $\lambda(x)_{\text{opt}} \cong p(x)$ for large k. Therefore, the smallest distortion attainable in the high rate regions is $D_{\text{opt}} = M(e^{H_x}/2^b)$, where H_x is the entropy of the stationary ergodic data source.

This result could have some applications to the design of large dimensional quantizers. Chen [8] and Linde et al. [9] consider the problem of choosing the output levels for multidimensional quantizers. In both papers, algorithms are presented to calculate the best choice of output levels assuming the input density is known. Linde et al. [9] also present a different algorithm which instead uses a set of training samples to calculate a "good" quantizer.

The above theory suggests, however, that at least in the large k case one does not really need to know the density function of the source. To design a quantizer for some stationary ergodic source one simply takes N k-dimensional samples from the source output and quantizes the succeeding k-dimensional data vectors by choosing from N samples the sample that minimizes the seminorm of the difference vector. This quantizer will be close to optimal (assuming large k) for any distortion measure satisfying our postulates.

IV. DISCUSSION

We should note that a key assumption in this derivation and in the derivations in [3] is that $p(x)$ may be assumed to be constant over small bounded sets. For the $L(x) = \|x\|_2^r$ distortion measure Zador [1] shows that it is sufficient that $p(x)$ be Riemann integrable. For arbitrary seminorms it is not clear exactly what the conditions on $p(x)$ must be.

We mention in the introduction that Zador derives an upper bound to the performance of block quantizers under the distortion $L(x) = \|x\|_2^r$. We should note that our D_u is looser than Zador's upper bound for this distortion because we have made use of the convexity properties of $M_k(v)$.

The rate throughout this paper has been assumed to be $k^{-1}\log N$. Some authors prefer to use the entropy of the quantizers output as the rate. From [2], [3], however, the quantizer output entropy is given by

$$H_Q \cong H_x - \frac{1}{k}\int p(y)\ln\frac{1}{N\lambda(y)}\,dy.$$

For large k the best output level point density is $\lambda_0(x) \cong p(x)$. Therefore, for the large dimensional optimal quantizer $H_Q \cong (1/k)\log N$.

Finally, we should note that at least for the cases where $(1/k)(D'_u - D_L) \to 0$ and $p(x)$ represents an ergodic source, we are calculating the distortion rate function for large rates. It is true that quantization theory is a subset of the source coding and rate distortion theory disciplines. Quantization theory is also in some senses more general than the present theory of source coding with respect to a fidelity criterion. Rate distortion theory as it has been developed to date has had limited success in setting performance bounds to anything except single letter distortion measures. The distortion measures considered in the paper (first put forward in [10] and generalized to multidimensions in [3]) are not under this constraint. Rate distortion theory sets performance bounds on source codes but gives few indications of how to achieve these bounds. On the other hand, the constructive nature of quantization theory gives deep insights into the implementation problem for these codes. Present day quantization results of course deal only with "nice" probability measures and asymptotic results. Relaxing these constraints must be a subject of further study. Until now rate distortion theory has provided the light to interpret vector quantizer performance. Perhaps in the future a synergistic effect between the two theories may begin to develop, greatly enriching both fields.

Devroye of McGill University for his proofs and helpful insights into p-varying functions.

REFERENCES

[1] P. Zador, "Development and evaluation of procedures for quantizing multivariate distributions," Ph.D. dissertation, Stanford University, Stanford, CA, University Microfilm 64-9855, 1964.

[2] A. Gersho, "Asymptotically optimal block quantization," *IEEE Trans. Inform. Theory*, vol IT-25, pp. 373–380, July 1979.

[3] Y. Yamada, S. Tazaki, and R. M. Gray, "Asymptotic performance of block quantizers with difference distortion measures," *IEEE Trans. Inform. Theory*, vol. IT-26, pp. 6–14, Jan. 1980.

[4] A. W. Roberts and D. E. Varberg, *Convex Functions*. New York: Academic, 1973.

[5] H. A. David, *Order Statistics*. New York: Wiley, 1970.

[6] A. Perez, "Generalization of Chernoff's result on the asymptotic discernibility of two random processes," *Colloquia Math. Soc. János Bolyai* (Budapest, Hungary), pp. 619–632, 1972.

[7] P. Billingsley, *Convergence of Probability Measures*. New York: Wiley, 1968.

[8] D. T. Chen, "On two or more dimensional optimum quantizers," in *Proc. 1977 IEEE Int. Conf. Acoustics, Speech, and Signal Processing*. 1977, pp. 640–643.

[9] Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantizer design," *IEEE Trans. Commun.*, vol. COM-28, pp. 84–95, Jan. 1980.

[10] H. Gish and J. N. Pierce, "Asymptotically efficient quantizing," *IEEE Trans. Inform. Theory*, vol. IT-14, pp. 676–683, Sept. 1968.

[11] W. Feller, *Introduction to Probability Theory and Its Applications*, vol. II. New York: Wiley, 1966.

A Dynamic Programming Algorithm for Phase Estimation and Data Decoding on Random Phase Channels

ODILE MACCHI, MEMBER, IEEE, AND LOUIS L. SCHARF, SENIOR MEMBER, IEEE

Abstract—The problem of simultaneously estimating phase and decoding data symbols from baseband data is posed. The phase sequence is assumed to be a random sequence on the circle, and the symbols are assumed to be equally likely symbols transmitted over a perfectly equalized channel. A dynamic programming algorithm (Viterbi algorithm) is derived for decoding a maximum *a posteriori* (MAP) phase-symbol sequence on a finite dimensional phase-symbol trellis. A new and interesting principle of optimality for simultaneously estimating phase and decoding phase-amplitude coded symbols leads to an efficient two-step decoding procedure for decoding phase-symbol sequences. Simulation results for binary, 8-ary phase shift keyed (PSK), and 16-quadrature amplitude shift keyed (QASK) symbol sets transmitted over random walk and sinusoidal jitter channels are presented and compared with results one may obtain with a decision-directed algorithm or with the binary Viterbi algorithm introduced by Ungerboeck. When phase fluctuations are severe and when occasional large phase fluctuations exist, MAP phase-symbol sequence decoding on circles is superior to Ungerboeck's technique, which in turn is superior to decision-directed techniques.

I. INTRODUCTION

PHASE FLUCTUATIONS can significantly increase the error probability for symbols transmitted over a channel that may or may not have been equalized. This is especially true for phase shift keyed (PSK) and quadrature amplitude shift keyed (QASK) symboling, in which case accurate phase discrimination is essential for symbol decoding. Even when the receiver contains a decision-directed phase-locked loop (DDPLL), performance loss in signal-to-noise ratio (SNR) with respect to a coherent decoding system can be in the range 5–10 dB. This fact is established in [1] for practical symbol sets and typical values of the phase variance parameter and symbol error probability.

On telephone lines, linear distortion and phase jitter dictate the use of a channel equalizer and some kind of phase estimator to achieve high rate, low error probability data transmission. A common approach to phase estimation and data decoding is to use a decision-directed algorithm in which a phase estimate is updated on the basis of old phase estimates and old symbol decisions. The DDPLL of [5] is a first-order digital phase-locked loop (PLL) in which the phase estimate is updated on the basis of a new measured phase and an old symbol decision. In the jitter equalizer (JE) of [3] and [4] a complex gain is updated according to a simple decision-directed stochastic ap-