Multivariate Analysis and Its Applications
IMS Lecture Notes – Monograph Series (1994) Volume 24

APPLICATIONS OF NUMBER-THEORETIC METHODS TO QUANTIZERS OF ELLIPTICALLY CONTOURED DISTRIBUTIONS

BY KAI-TAI FANG, KE-HAI YUAN AND PETER M. BENTLER*

Hong Kong Baptist College, University of California, Los Angeles, and University of California, Los Angeles

In this paper we combine the so-called LBG algorithm in vector quantization and the number-theoretic methods to propose a new algorithm, NTLBG, of generating vector quantizer with low mean square error for many classes of multivariate distributions including one of the elliptically contoured distributions. Some numerical examples show that the present method is effective.

1. Introduction. The number-theoretic (or quasi Monte Carlo) method has been applied in many branches of statistics: numerical evaluation of probabilities and moments of multivariate distributions, optimization, regression analysis (nonlinear and robust regression, and regression with constraints), estimation and testing hypothesis, experimental design and geometric probability, see Wang and Fang (1981,1990a,1990b,1991), Shaw (1988), Fang, Yuan and Bentler (1992), Fang and Wang (1991), Fang, Zhu, and Bentler (1993)). Fang and Wang (1993) and Fang, Wang and Bentler (1994) gave a comprehensive review in a bibliographic setting. These works show that the number-theoretic method (NTM) is useful and powerful in many statistical problems. In this paper, we will apply the NTM to quantization and give a unified approach to generating a quantizer of an elliptically contoured distribution with low mean square error.

Let X be a random variable with p.d.f. $p(x)$ and $\sigma^2 = \mathrm{Var}(X)$. For a given integer n we want to find $-\infty < x_1 < x_2 < \cdots < x_n < \infty$ such that the mean square error (MSE)

---

* This paper is a small part of our presentation at The International Symposium on Multivariate Analysis and Its Applications. The primary contents of our presentation will be published in Statistical Science this year.

AMS 1991 Subject Classifications: 94A34, 62H99

Key words and phrases: Elliptically contoured distribution, F-discrepancy, mean square error, number-theoretic method, quantizer, quasirandom number.

$$\text{MSE}(x_1, \cdots, x_n) = \frac{1}{\sigma^2} \int_{-\infty}^{\infty} \min_{1 \le i \le n} (x - x_i)^2 p(x)\, dx \tag{1.1}$$

is minimum over $TR_n = \{(x_1, \cdots, x_n) : x_1 < x_2 < \cdots < x_n\}$. The corresponding $\{x_i\}$ are called rep-points (representative points) of X that can be definded in another way: For any $(x_1, \cdots, x_n) \in TR_n$, its quantizer $Q_n(.)$ is definded as the step function:

$$Q_n(x) = x_k, \qquad \text{if } a_k < x \le a_{k+1}, k = 1, \cdots, n \tag{1.2}$$

where

$$a_1 = -\infty, \quad a_{n+1} = \infty, \quad a_k = (x_{k-1} + x_k)/2, \quad k = 2, \cdots, n.$$

Then the MSE of $\{x_i\}$ or $Q_n(x)$ defined in (1.2) can be expressed as

$$\text{MSE}(Q) = \text{MSE}(x_1, \cdots, x_n) = \frac{1}{\sigma^2} \text{E}(X - Q_n(X))^2. \tag{1.3}$$

The optimum quantizer $Q_n^*(x)$ minimizes $\text{MSE}(x_1, \cdots, \boldsymbol{x}_n)$ over TR_n. The corresponding solution is just the rep-points of X. The problem for finding a set of rep-points has been raised in many fields, such as information theory, cluster analysis, correlation analysis, signal processing, theory of quantization, and dress standardization (Cox (1957), Bofinger (1970), and Fang and He (1984)). Max (1960), Lloyd (1982), Fang and He (1984), and Cambanis and Gerr (1983) have proposed using numerical methods for finding rep-points of a given distribution. With the sequential number-theoretic method for optimization (SNTO) suggested by Wang and Fang (1990b) we can find approximate rep-points for any continuous distribution with finite variance (Fang and Wang (1991)).

In this paper we shall concentrate our study on quantizers of a multivariate distribution. Let $\boldsymbol{x} = (X_1, \cdots, X_s)'$ be an s-dimensional random vector with joint density $p(x_1, \cdots, x_s)$. An n-level quantizer $Q = \{\mathcal{Y}, \mathcal{S}\}$ of $\boldsymbol{x}$ consists of

(1) a set of output vectors $\mathcal{Y} = \{\boldsymbol{y}_1, \cdots, \boldsymbol{y}_n\}$;

(2) a partition $\mathcal{S} = (S_1, \cdots, S_n)$ of the space R^s with n disjoint and exhaustive regions;

(3) a mapping $Q : R^s \to \mathcal{S}$ defined by $Q(\boldsymbol{x}) = \boldsymbol{y}_i$, if $\boldsymbol{x} \in S_i$.

When $s > 1$ people often call the quantizer a block quantizer, vector quantizer, or s-dimensional quantizer. The optimal n-level quantizer minimizes

$$\text{MSE}(Q) = \frac{1}{s} \text{E} \|\boldsymbol{x} - Q(\boldsymbol{x})\|, \tag{1.4}$$

the mean square error, over all n-level quantizers and the corresponding $\{\boldsymbol{y}_i\}$ are called rep-points of $\boldsymbol{x}$.

It is easy to find the following necessary conditions for an optimal quantizer:

alignment 1) associated with each quantizer output $\boldsymbol{y}_i$, there is a partition cell S_i, $i = 1, \cdots, n$, satisfying

$$S_i = \{\boldsymbol{x} : \|\boldsymbol{x} - \boldsymbol{y}_i\| < \|\boldsymbol{x} - \boldsymbol{y}_j\|, j \neq i\}; \tag{1.5}$$

alignment 2) for each patition cell $S_i, the \boldsymbol{y}_i$ should equal the conditional mean

$$\boldsymbol{y}_i = \mathrm{E}[\boldsymbol{x}|\boldsymbol{x} \in S_i]. \tag{1.6}$$

When $s > 1$ it is difficult to find the optimum quantizer as well as the rep-points for any multivariate distribution. A theoretical basis for asymptotically optimum vector quantization has been provided by Zador (1982), and Gersho (1979, 1982). Zador found for the asymptotic case of high quality quantization that

$$\mathrm{MSE}(Q) = A(s,2)n^{-2/s}\|p(\boldsymbol{x})\|_{s/(s+r)} \tag{1.7}$$

where

$$\|p(\boldsymbol{x})\|_t = [\int |p(\boldsymbol{x})|^t d\boldsymbol{x}]^{1/t}$$

is the l_t-norm and $A(s,2)$ is independent of the density $p(\boldsymbol{x})$. When n is not large, Linde, Buzo and Gray (1980) suggested an iterative vector quantizer algorithm LBG based on a training sequence. This approach has been used to calculate low bit-rate quantizers for multivariate normal, multivariate Laplace, multivariate gamma and multivariate uniform distributions (see, for example, Abut, Gray, and Rebolledo (1982), Gray and Linde (1982), and Gray (1984)).

The necessary conditions (1.5) and (1.6) provide the basis of the LBG algorithm, the details of which are given in Linde et. al. (1980). The outline of the algorithm can be stated as follows:

1) Set $t = 0$. Give an initial set of output vectors $\mathcal{Y}_t = \{\boldsymbol{y}_{t1}, \cdots, \boldsymbol{y}_{tn}\}$ and find the associated partion $\mathcal{S}_t = (S_{t1}, \cdots, S_{tn})$ by (1.5).

2) Generating a training sequence $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_N$, by the Monte Carlo method, representing the distribution of random vector $\boldsymbol{x}$.

3) Each $\boldsymbol{x}_i$ is assigned to the nearest sell of the partion $\mathcal{S}_t$.

4) Calculate the sample mean, $\boldsymbol{y}_{t+1,j}$ say, of $\boldsymbol{x}_i$ which falling in the cell $S_{t,j}, j = 1, \cdots, n$. Let $\mathcal{Y}_{t+1} = \{\boldsymbol{y}_{t+1,1}, \cdots, \boldsymbol{y}_{t+1,n}\}$. If $\mathcal{Y}_{t+1} = \mathcal{Y}_t$, then output $\mathcal{Y}_t$ as the final output vector, otherwise go to next step.

5) Let $t = t + 1$ and go to step 3).

This process provides a nonincreasing MSE, and the algorithm eventually converges to a locally optimal solution $\mathcal{Y} = \{\boldsymbol{y}_1, \cdots, \boldsymbol{y}_n\}$.

The LBG algorithm is essentially based on the k-means method of cluster analysis. It is interesting to note that there are so many authors (for example, MacQueen (1967), Pollard (1981, 1982), Gray and Karnin (1982), and Zhang and Fang (1982)) who have given different proofs for convergence of the k-means algorithm.

The LBG algorithm has the overwhelming advantage of being a general algorithm that can generate a vector quantizer. But there are some problems with the approach:

1) The resulting output vectors are only locally optimum and depend on the initial set of output vectors. There is no unified way to offer a good initial set of output vectors so far.

2) The training sequence and numerical evaluation of MSE has been based on the Monte Carlo method. The convergence rate of Monte Carlo method is $O(n^{-1/2})$ in probability and is slow.

There are a number of authors who have proposed many kinds of the initial output vectors. Wilson (1980), Fisher and Dicharry (1984), Fisher (1989) gave geometric designs. An alternative approach is the lattice quantizer by Conway and Sloane (1983), Gersho (1979) and Sayood et. al. (1984), in which they considered only the uniform distribution. Fisher (1986) proposed a source coding approach based on Shannon's entropy theory (1948). This approach is effective only for very high dimensional quantizers. Already there is work done in this direction; for example, Fisher (1986), Gersho et. al. (1983), Sabin and Gray (1984) and Tseng and Fisher (1987). However, one wishes to find a unified way to offer initial output vector which can produce a vector quantizer with low MSE value by the LBG algorithm for various multivariate distributions. In this paper you will see that the use of NTM can meet this requirement well.

The essence of NTM is to find a set of points that is uniformly scattered over an s-dimensional unit cube $[0,1]^s$ (it is called a set of quasirandom numbers or an NT-net on $[0,1]^s$) and sometimes this set can be used instead of random numbers in the Monte Carlo method with better results. Therefore the NTM is also known as the quasi or deterministic version of the Monte Carlo method. The application of NTM is a new, but rapidly expanding, branch of statistics. In this paper the NTM is applied to give a new algorithm, NTLBG, for generating low MSE quantizer and the associated feasible rep-points of elliptically contoured distributions (ECD) including multivariate normal, multivariate t-, multivariate uniform distributions and related distributions. The algorithm modifies the LBG method in the following aspect:*the initial set of output vectors as well as the training sequence are generated by the NTM.* Our numerical results show that the present method can be used to obtain low MSE quantizers for ECD and various classes of distributions.

The paper is organized as follows. Section 2 develops a unified approach to generating quasirandom F-sequences for many classes of multivariate dis-

tributions including one of ECD. A new algorithm NTLBG of generating vector quantizer is given in Section 3. Some numerical results for multivariate normal, symmetric multivariate Pearson Type VII and multivariate uniform distributions are given in Section 4.

2. Quasirandom F-numbers. The number-theoretic methods are based on the so-called quasirandom F-numbers which will be defined below. First we need the concept of F-discrepancy proposed by Wang and Fang (1990a).

DEFINITION 1. Let $\mathcal{P} = \{\boldsymbol{x}_k, k = 1, \cdots, n\}$ be a set of points in R^s and $F_n(\boldsymbol{x})$ be its empirical distribution, i.e.,

$$F_n(\boldsymbol{x}) = \frac{1}{n}\sum_{i=1}^{n} I\{\boldsymbol{x}_i \leq \boldsymbol{x}\},$$

where $I\{A\}$ is the indicator function of A, and all inequalities are understood with respect to the componentwise order of R^s. Then

$$D_F(n, \mathcal{P}) = \sup_{\boldsymbol{x} \in R^s} |F_n(\boldsymbol{x}) - F(\boldsymbol{x})| \tag{2.1}$$

is said to be the F-discrepancy of $\mathcal{P}$ with respect to $F(\boldsymbol{x})$.

When $F(\boldsymbol{x})$ is the uniform distribution $U(C^s)$ the F-discrepancy reduces to common discrepancy in the sense of Weyl (1916). In fact, the F-discrepancy is just the Kolmogorov-Smirnov distance in goodness-of-fit tests.

DEFINITION 2. A set of points $\mathcal{P}_n^* = \{\boldsymbol{p}_k^*, k = 1, \cdots, n\}$ in R^s is called the F-optimum set of points with respect to $F(\boldsymbol{x})$ if

$$D_F(\mathcal{P}_n^*) = \inf_{\mathcal{P}_n} D_F(\mathcal{P}). \tag{2.2}$$

where $\mathcal{P}_n = \{\boldsymbol{P}_k : \boldsymbol{P}_k \in R^s, k = 1, \cdots, n\}$.

The following facts show that there always exists the F-optimum set of points for each univariate continuous distribution.

1) The set $Q_n = \{(2i-1)/(2n), i = 1, \cdots, n\}$ is the optimum set of points with respect to the uniform distribution $U([0,1])$ and with discrepancy 1/2n.

2) Let $F(x)$ be a continuous distribution function and let $F^{-1}(x)$ be its inverse function. Then the set $Q_n^F = \{F^{-1}((2i-1)/(2n)), i = 1, \cdots, n\}$ is the F-optimum set of points with respect to $F(x)$ and with F-discrepancy 1/2n.

So far it is difficult to find the F-optimal set of points for multivariate distributions. So we have to be content with the second best.

DEFINITION 3. Let $\mathcal{P}_n$ be a sequence of sets of points under a certain structure in R^s and $F(\boldsymbol{x})$ be an s-dimensional distribution function. If $D_F(\mathcal{P}_n) = O(n^{-1}(\log n)^s)$ as $n \to \infty$, the points of $\mathcal{P}_n$ are called quasirandom F-numbers, and the $\mathcal{P}_n$ is called quasirandom F-numbers sequence.

When $F(\boldsymbol{x})$ is the uniform distribution $U(C^s)$, where $C^s = [0,1]^s$ is a unit cube, the quasirandom F-numbers reduce to the common quasirandom numbers. There are many methods, such as the good lattice point (*glp*) method, the good point (*gp*) method, the Halton method and (t,s)-sequence, that can be used to produce quasirandom number sequences in C^s (cf. Hua and Wang (1981), Niederreiter (1992)).

There is no universally applicable method for generating quasirandom F-numbers. fortunately. The following method can be applied to most useful multivariate distributions.

Let $F(\boldsymbol{x})$ be a continuous distribution in R^s and $\boldsymbol{x}$ be a random vector such as $\boldsymbol{x} \sim F(\boldsymbol{x})$. Suppose that $\boldsymbol{x}$ has a stochastic representation

$$\boldsymbol{x} = \boldsymbol{h}(\boldsymbol{y}), \tag{2.3}$$

where $\boldsymbol{y} \sim U(C^t), t \leq s$ and $\boldsymbol{h}$ is a continuous function on C^t. Let $\{\boldsymbol{c}_k, k = 1, \cdots, n\}$ be a set of quasirandom numbers on C^t and let $\boldsymbol{x}_k = \boldsymbol{h}(\boldsymbol{c}_k), k = 1, \cdots, n$. Then $\{\boldsymbol{x}_k\}$ is a set of quasirandom F-numbers. Wang and Fang (1990a) pointed out that the above method has good properties. In particular, the method can be applied to generating quasirandom F-numbers of elliptically contoured and multivariate Liouville distributions(cf. Fang, Kotz and Ng (1990)).

An s-dimensional random vector $\boldsymbol{x}$ is said to have a spherically symmetric distribution if $\boldsymbol{x}$ and $\boldsymbol{P}\boldsymbol{x}$ have the same distribution for every orthogonal matrix $\boldsymbol{P}$. It is known that $\boldsymbol{x}$ has the stochastic representation (SR)

$$\boldsymbol{x} = R\boldsymbol{u}^{(s)}, \tag{2.4}$$

where random variable $R \geq 0$ is independent of $\boldsymbol{u}^{(s)}$ which is uniformly distributed on the unit sphere, U_s say. Wang and Fang (1990a) proposed an algorithm for generating quasirandom F-numbers of the uniform distribution on U_s by the use of a spherical coordinate transformation. Another more efficient algorithm (the so-called TFWW algorithm) is due to Tashiro (1977) and Fang, Wang and Wong (1992). For details of the TFWW algorithm see Fang and Wang (1993).

Combining the TFWW algorithm and (2.4) we obtain the so-called NTSR algorithm:

Step 1. Generate a set of quasirandom numbers $\{\boldsymbol{c}_k = (c_{k1}, \cdots, c_{ks}), k = 1, \cdots, n\}$.

Step 2. Denote the cdf of R by $F_R(r)$ and let ${F_R}^{-1}$ be its inverse function. Compute $r_k = {F_R}^{-1}(c_{ks}), k = 1, \cdots, n$.

Step 3. Generate a set of quasirandom F-numbers $\{\boldsymbol{y}_k, k = 1, \cdots, n\}$ of the uniform distribution on U_s with the first (s-1)-components of $\boldsymbol{c}_k's$.

Step 4. Then $\{\boldsymbol{x}_k = r_k\boldsymbol{y}_k, k = 1, \cdots, n\}$ is a set of quasirandom F-numbers of the given spherical distribution $F(\boldsymbol{x})$.

With this algorithm we can generate quasirandom F-number sequence for every continuous spherical distribution (in this case R defined in (2.4) is a continuous distribution). Obviously, the NTSR algorithm can be similarly applied to many kinds of distributions such as the multivariate l_1-norm symmetric and multivaraite Liouville distributions.

An s-dimensional random vector $\boldsymbol{x}$ is said to have an elliptically contour distribution with parameters $\boldsymbol{\mu}(s \times 1)$ and $\boldsymbol{\Sigma}(s \times s)$ if

$$\boldsymbol{x} = \boldsymbol{\mu} + \boldsymbol{A}'\boldsymbol{y}, \tag{2.5}$$

where $\boldsymbol{y}$ has a spherically symmetric distribution and $\boldsymbol{A}'\boldsymbol{A} = \boldsymbol{\Sigma}$. Obviously, we can generate quasirandom F-numbers for every continuous elliptically contoured distribution.

3. An Algorithm For Vector Quantizers. Based on the LBG algorithm and the NTM we would propose an algorithm NTLBG for vector quantizers of a multivariate distribution $F(\boldsymbol{x})$ with a stochastic representation (2.3). The NTLBG algorithm involves the following steps:

Step 1. For given $F(\boldsymbol{x})$ generate a set of quasirandom F-numbers $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_N$ as a training sequence by the NTM mentioned in section 2. The number N is often large.

Step 2. Set $t = 0$. For given n generate a set of quasirandom F-numbers $\boldsymbol{y}_{t1}, \cdots, \boldsymbol{y}_{tn}$ of $F(\boldsymbol{x})$ as an initial set of output vectors $\mathcal{Y}_t = \{\boldsymbol{y}_{t1}, \cdots, \boldsymbol{y}_{tn}\}$.

Step 3. Form a partition $\{S_i\}$ of $\{\boldsymbol{x}_k, k = 1, \cdots, N\}$ such that each $\boldsymbol{x}_i$ is assigned to the nearest cell of a partition, i.e. $\boldsymbol{x}_i \in S_j$ if $\|\boldsymbol{x}_i - \boldsymbol{y}_{tj}\| \leq \|\boldsymbol{x}_i - \boldsymbol{y}_{tk}\|, k \neq j$ (cf.(1.5)).

Step 4. Calculate the sample conditional means (1.6) and from a new set of output vector vectors $\mathcal{Y}_{t+1} = \{\boldsymbol{y}_{t+1,1}, \cdots, \boldsymbol{y}_{t+1,n}\}$, where

$$\boldsymbol{y}_{t+1,j} = \frac{1}{n_{t+1,j}} \sum_{k \in S_j} \boldsymbol{x}_k,$$

and $n_{t+1,j}$ is the number of $\boldsymbol{x}_k$ falling in S_j.

Step 5. If $\mathcal{Y}_{t+1} = \mathcal{Y}_t$ terminate the algorithm and deliver the final output $\mathcal{Y}_t$. Otherwise let $t = t + 1$ and return to step 3.

We now give some explanations for the above steps.

Step 2): People often use the F-discrepancy for measuring uniformity of quasirandom F-numbers, but use the MSE (or other error distortion measure) for output vectors of a quantizer. These two measures are different, but there are some relationships between them. For example, take a sequence of sets of $\mathcal{P}_n$ of points in C^s with discrepancy $D(\mathcal{P}_n) = O(n^{-1}\log^s n)$. Then we have

$$\max_{\boldsymbol{x}\in C^s} \min_{1\le k\le n} \|\boldsymbol{x} - \boldsymbol{x}_k\| \le O(n^{-1/s}\log n), \tag{3.1}$$

where $\|\boldsymbol{x} - \boldsymbol{x}_k\|$ is the l_2-norm (cf. Wang and Fang (1990a)) and

$$\text{MSE} = \frac{1}{s}\int_{C^s} \min_{1\le k\le n} \|\boldsymbol{x} - \boldsymbol{x}_k\|^2 \, d\boldsymbol{x} \le O(n^{-2/s}\log^2 n) \tag{3.2}$$

which is close to the order $O(n^{-2/s})$ in (1.7) as $\|p(\boldsymbol{x})\| = 1$. The results in the next section show that the initial points chosen by this way can obtain a vector quantizer with low MSE in most cases. Thus, we provide a unified way to choose an initial output for every multivariate distribution which has a stochastic representation (2.3).

Step 1): A training sequence generated by NTM has a better representation than one generated by Monte Carlo methods. Furthermore, the symmetrized technique (or the antithetic variates technique in Monte Carlo methods) is widely used in numerical integration and is more efficient (cf. Zaremba(1972)). For example, if $\{\boldsymbol{c}_k = (c_{k1}, c_{k2})\}$ is a set of points with low discrepancy, its symmetrization set is given by

$$\{(c_{k1}, c_{k2}), (c_{k1}, 1-c_{k2}), (1-c_{k1}, c_{k2}), (1-c_{k1}, 1-c_{k2}), k = 1, \cdots, N\}$$

which has $4N$ points, and the corresponding quasirandom F-numbers follow. Our results (not put into this work) show that this kind of training sequence can save much computing time and make higher precision of MSE.

4. Quantizer of Elliptically Contoured Distributions. In this section we apply the algorithm proposed in the previous section to obtain low MSE quantizers for some elliptical distributions and give comparisons among our results and other results. Throughout this section we always employ the *glp* method to generate sets of quasirandom numbers. For given n and s choose an integer vector $\boldsymbol{h} = (h_1, \cdots, h_s)$ satisfying $1 \le h_i < n, h_i \ne h_j$ for $i \ne j$. Calculate

$$c_{ki} = \{(2kh_i - 1)/2n\}, k = 1, \cdots, n; i = 1, \cdots, s,$$

where $\{x\}$ denotes the fraction part of x. The choice of $\boldsymbol{h}$ can be found in the Appendix of Hua and Wang (1981) or Fang and Wang (1993). Then $\boldsymbol{c}_k = (c_{k1}, \cdots, c_{ks}), k = 1, \cdots, n$ is a *glp* set of quasirandom numbers.

EXAMPLE 1. The Multivariate Normal Distribution.

Without loss of generality we consider the standard normal $N_s(\mathbf{0}, \boldsymbol{I}_s)$ only. Let $\boldsymbol{x} \sim N_s(\mathbf{0}, \boldsymbol{I}_s)$ with the cdf $\Phi(x_1, \cdots, \boldsymbol{x}_s)$. It is known that $\Phi(x_1, \cdots, x_s) = \Phi(x_1) \cdots \Phi(x_s)$ where $\Phi(x)$ is the cdf of $N(0,1)$. Let $\{\boldsymbol{c}_k = (c_{k1}, \cdots, c_{ks}), k = 1, \cdots, n\}$ be a set of quasirandom numbers by the *glp* method. Then $\{\boldsymbol{x}_k = (x_{k1}, \cdots, x_{ks}), k = 1, \cdots, n\}$, where $x_{kj} = \Phi^{-1}(c_{kj})$, is a set of quasirandom F-numbers for $N_s(\mathbf{0}, \boldsymbol{I}_s)$. By this way we can generate a training sequence $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_N$ as well as an initial set of output vector. For illustration we consider only the binormal with $N = 70844$ and $s = 3 \sim 32$. By the use of NTLBG algorithm the corresponding quantizers are showing in Figue 1 in the Appendix of the paper. The MSE values are plotted (with "×") in Figure 1 and listed in Table 1. We can see that our results are better than Wilson's (1980) and Fisher and Dicharry's(1984), the latter gave quantizers only for $n = 8, 16, 32$ by their special geometric design.

Table 1

n	MSE	n	MSE	n	MSE
3	0.4618390	13	0.1301080	23	0.0771025
4	0.3623600	14	0.1227065	24	0.0744985
5	0.3063070	15	0.1172445	25	0.0714630
6	0.2711950	16	0.1067880	26	0.0689305
7	0.2216000	17	0.1020085	27	0.0659590
8	0.2002523	18	0.0967775	28	0.0639525
9	0.1806665	19	0.0922800	29	0.0620685
10	0.1635840	20	0.0875885	30	0.0600280
11	0.1513470	21	0.0835925	31	0.0582365
12	0.1387165	22	0.0800835	32	0.0569415

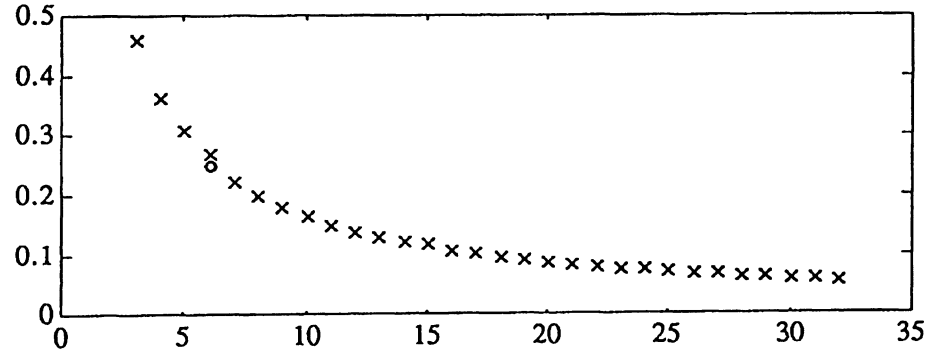

Figure 1 MSE values of quantizers of normal distribution

Since the normal distribution $N_s(\mathbf{0}, \boldsymbol{I}_s)$ is a spherical distribution with χ_s in (2.4), where χ_s is the Chi-distribution with s degrees of freedom, we might use the NTSR algorithm mentioned in Section 2 to generate $\boldsymbol{u}^{(s)}$ in (2.4) by the TFWW algorithm and the inverse transform method to generate χ_s. Consequently, we can find a set of quasirandom F-numbers for $N_s(\mathbf{0}, \boldsymbol{I}_s)$. When $s = 2$, the output vectors by this method are very close to that of the first method except for an orthogonal transformation. For example, for $n = 8, 16, 32$, the MSE values are 0.200268, 0.106794 and 0.056556 respectively, which are at the same level as in Table 1.

REMARK: Wang and Fang (1981) proposed a method to obtain a set of quasirandom numbers for even n with lower discrepancy by deleting the last point of a *glp* set of size $n + 1$ of quasirandom numbers and rescaling these numbers. Their method provides another way to design the initial set of output vectors. For example, when $n = 6$, by this way the output vectors with MSE = 0.2522955 is plotted by 'o' in Figure 1, which is better than that with MSE = 0.2711950 in Figure 1. Their output vectors are plotted in Figure 2(b) and the output vectors obtained before are plotted in Figure 2(a). Similarly we can obtain a set of output vectors in the case of $n = 32$ with MSE = 0.056895 which gives a slightly improved results than MSE = 0.0569415 in Table 1.

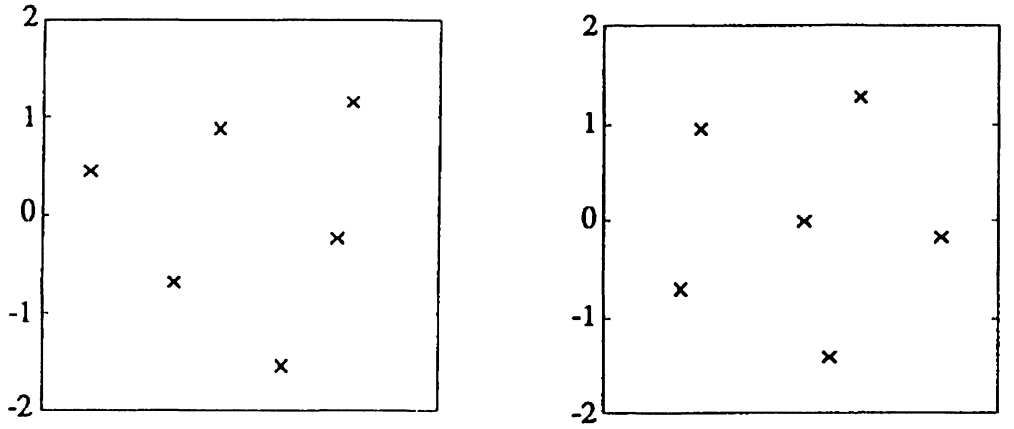

Figure 2

EXAMPLE 2. Some Spherical Distributions.

Let $\boldsymbol{x}$ have a symmetric multivariate Pearson Type VII distribution (cf. Fang, Kotz and Ng (1990), 81–82). The density function of the associated R (cf. (2.4)) is given by

$$\frac{2}{B(s/2, N - s/2)} m^{-s/2} r^{s-1} (1 + r^2/m)^{-N}.$$

It is easy to generate quasirandom F-numbers of $\boldsymbol{x}$ by the NTSR algorithm mentioned in section 2. When $m = 4, N = 15$, and $s = 2$ the NTLBG produces sets of theoutput vectors for $n = 8, 16, 32$ with respective MSE values $0.032294, 0.017586, 0.009286$.

When $\boldsymbol{x}$ is uniformly distributed in the unit ball, $\boldsymbol{x}$ has a spherical distribution. The corresponding density function of R in (2.4) is sr^{s-1}, $0 \leq r \leq 1$, (cf. Fang, Kotz and Ng (1990),74–75). Its quasirandom uniform-numbers can be generated similarly. The output vectors in the case of $n = 8, 16, 32$ have MSE values $0.032936, 0.016591, 0.008116$, respectively.

The above two examples indicate that the new algorithm is effective to generate sets of output vectors for all continuously spherical and elliptical distributions as well as for many classes of multivariate distributions, such as multivariate Liouville distributions.

We also applied the NTLBG to the Laplacian and memoryless gamma distributions and found that most of our results are better than Fisher and Dicharry's (1984).

Acknowledgment: This work was supported by Honk Kong UPGC grant and PHS National Institute on Drug Abuse Award DA01070-20. The authors would like to thank Prof. T. R. Fisher and Prof. S. Cambanis for kindly supplying copies of relevant papers. We would also like to thank Professor Y. L. Tong and Dr. K. W. Ng for valuable comments.

REFERENCES

ABUT, H., GRAY, R. M. and REBOLLEDO, G. (1982). Vector quantization of speech and speech-like waveforms, *IEEE Trans. Acoust., Speech, Signal Processing*, bf ASSP-30, 423–435.

BOFINGER, E. (1970). Maximizing the correlation of grouped observations, *J. Amer. Statist. Assoc.*, **65**,1632–8.

CAMBANIS, S. and GERR, N. L. (1983). A simple class of asymptotically optimal quantizer, *IEEE Trans. Inform. Theory*, **IT-29**, 664–676.

CONWAY, J.H. and SLOANE, N.J.A. (1983). A fast encoding method for lattice codes and quantizers, *IEEE Trans. Inform. Theory*, **IT-29**, 820–824.

COX, D.R. (1957). Note on grouping, *J. Amer. Statist. Assoc.*, **65**, 543–547.

FANG, K. T. and HE, S.D. (1984). The problem of selecting a specified number of representative points from a normal population, *Acta Math. Appl. Sinica*, **7**, 293–306.

FANG, K. T., KOTZ, S.,and NG, K. W. (1990). *Symmetric Multivariate and Related Distributions*, Chapman and Hall, London/New York.

FANG, K. T.and WANG, Y. (1991). A sequential algorithm for solving a system of nonlinear equations, *J. Comput. Math.*, **9**, 9–16.

FANG, K.T. and WANG, Y. (1993). *Number-theoretic Methods in Statistics*, Chapman and Hall, London/New York.

FANG, K. T., WANG, Y. and BENTLER, P. M. (1994). Some applications of number-theoretic methods in statistics, *Statistical Science*, to appear.

FANG, K.T., WANG, Y. and WONG, H. L. (1992). A new method of generating an NT-net on the unit sphere, Technical report MATH-002, Hong Kong Baptist College, Hong Kong.

FANG, K. T. and YUAN, K. H. (1990). A unified approach to maximum likelihood estimation, *Chinese J. Appl. Prob. Stat.*, **6**, 412–418.

FANG, K. T., YUAN, K. H. and BENTLER, P. M. (1992). Applications of sets of points uniformly distributed on sphere to testing multinormality and robust estimation, in *"Probability and Statistics"*, eds. by Jiang, Z.P. et. al., World Scientific, 56–73.

FANG, K. T., ZHU, L. X. and BENTLER, P. M. (1993). A necessary test of goodness of fit for sphericity, *J. Multivariate Analysis*, **45**, 34–55.

FISHER, T. R. (1986). A pyramid vector quantizer, *IEEE Trans. Inform. Theory*, **IT-32**, 568–583.

FISHER, T. R. (1989). Geometric source coding and vector quantization, *IEEE Trans. Inform. Theory*, **IT-35**, 137–145.

FISHER, T. R. and DICHARRY, R. M. (1984). Vector quantizer design for memoryless Gaussian, gamma, and Laplacian sources, *IEEE Trans. Commun.*, **COM-32**, 1065–1069.

GERSHO, A. (1979). Asymptotically optimal block quantization, *IEEE Trans. Inform. Theory*, **IT-25**, 373–380.

GERSHO, A. (1982). On the structure of vector quantizers, *IEEE Trans. Inform. Theory*, **IT-28**, 157–166.

GERSHO, A. and CUPERMAN, V. (1983). Vector quantization: A pattern-matching technique for speech coding, *IEEE Commun. Mag.*, **21**, 15–21.

GREY, R. M. (1984). Vector quantization, *IEEE ASSP Mag.*, **1**, 4–29.

GREY, R. M. and KARNIN, E. D. (1982). Multiple local optima in vector quantizers, *IEEE Trans. Inform. Theory*, **IT-28**, 256–261.

GREY, R. M. and LINDE, Y. (1982). Vector quantizers for Gauss-Markov sources, *IEEE Trans. Commun.*, **COM-30**, 381–389.

HUA, L. K. and WANG, Y. (1981). *Application of Number Theory to Numerical Analysis*, Spring-verlag and Science Press, Berlin/Beijing.

LINDE, Y. L., BUZO, A. and GRAY, R. M. (1980). An algorithm for vector quantizer design, *IEEE Trans. Commun.*, **COM-28**, 84–95.

LLOYD, S. P. (1982). Least squares quantization in PCM, *IEEE Trans. Inform. Theory*, **IT-28**, 129–137.

MACQUEEN, J. (1967). Some method for classification and analysis of multivariate observations, in *Proceeding 5th Berkeley Symposium on Math. Statist. and Prob.*, **Vol. 1**, 281–296.

MALONE, K. T. and FISHER, T. R. (1988). Contour-gain vector quantization, *IEEE Trans. Acoust., Speech, Signal Processing*, **ASSP-36**, 862–870.

MAX, J. (1960). Quantizing for minimum distortion, *IRE Trans. theory*, **IT-6**, 7–12.

NIEDERREITER, H. (1992). *Random Number Generation and Quasi-Monte Carlo methods*, SIAM, Philadelphia.

POLLAD, D. (1981). Strong consistency of k-means clustering, *Ann. Statist.*, **9**, 135–140.

POLLAD, D. (1982). Quantization and the method of k-means, *IEEE Trans. Inform. Theory*, **IT-28**, 199–205.

SABIN, M. J. and GRAY, R. M. (1984). product code vector quantizers for waveform and voice coding, *IEEE Trans. Acoust., Speech, Signal Processing* , **ASSP-32**, 474–488.

SAYOOD, K., GIBSON, J. D. and ROST, M. C. (1984). An algorithm for uniform vector quantizer design, *IEEE Trans. Inform. Theory*, **IT-30**, 805–814.

SHANNON, C. E. (1948). A mathematical theory of communication, *Bell Syst. Tech. J.*, **27**, 379–423 and 623–656.

SHAW, J. E. H. (1988). A quasirandom approach to integration in Bayesian statistics, The Ann. Statist.,**16**,859–914.

TASHIRO, Y. (1977). On methods for generating uniform points on the surface of a sphere, *Annals of Institute of Statistical Mathematics*, **29**, 295–300.

TSENG, H. C. and FISHER, T. R. (1987). Transform and hybrid transform/ DPCM coding of images using pyramid vector quantization, *IEEE Trans. Commun.*, **COM-35**, 79–86.

WANG, Y.and FANG, K. T. (1981). A note on uniform distribution and experimental design, *Kexue Tongbao*, **26**, 485–496.

WANG, Y. and FANG, K. T. (1990a). Number-theoretic method in applied statistics, *Chinese Ann. Math. Ser. B*, **11**, 31-55.

WANG, Y. and FANG, K. T. (1990b). Number-theoretic method in applied statistics, *Chinese Ann. Math. Ser. B*, **11**, 76–86.

WANG, Y. and FANG, K. T. (1991). A sequential number-theoretic method for optimization and its applications in statistics, in *Statistics in China: Achievements and Progress*, Longman, London, 139–156.

WEYL, H. (1916). Uber die Gleichversteilung der Zahlen mod, *Eins, Math. Ann.*, **77**, 313–352.

WILSON, S. G. (1980). Magnitude/Phase quantization of independent Gaussian variates, *IEEE Trans. Commun.*, **COM-28**, 1924–1929.

ZADO, P.L. (1982). Asymptotic quantization error of continuous signals and quantization dimension, *IEEE Trans. Inform. Theory*, **IT-28**, 139–149.

ZAREMBA, S. K. (1972). Le Méthod des "bons treillis"pour le caleul des intégrales, in *Applications of Number Theory to Numerical Analysis*, (ed. by Zaremba). Academic Press, New York.

ZHANG, Y. T. and FANG, K. T. (1982). *An Introduction to Multivariate Analysis*, Science Press, Beijing.

ZHU, L. X., WONG, H. L. and FANG, K. T. (1993). A test for multinormality based on sample entropy and projection pursuit, *J. Statist. Planning and Inference*, to appear.

DEPARTMENT OF MATHEMATICS
HONG KONG BAPTIST COLLEGE,
KOWLOON TONG, KOWLOON
HONG KONG

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA
LOS ANGELES CA 90024–1555
USA

DEPARTMENT OF PSYCHOLOGY
UNIVERSITY OF CALIFORNIA
LOS ANGELES CA 90024–1563
USA

Appendix

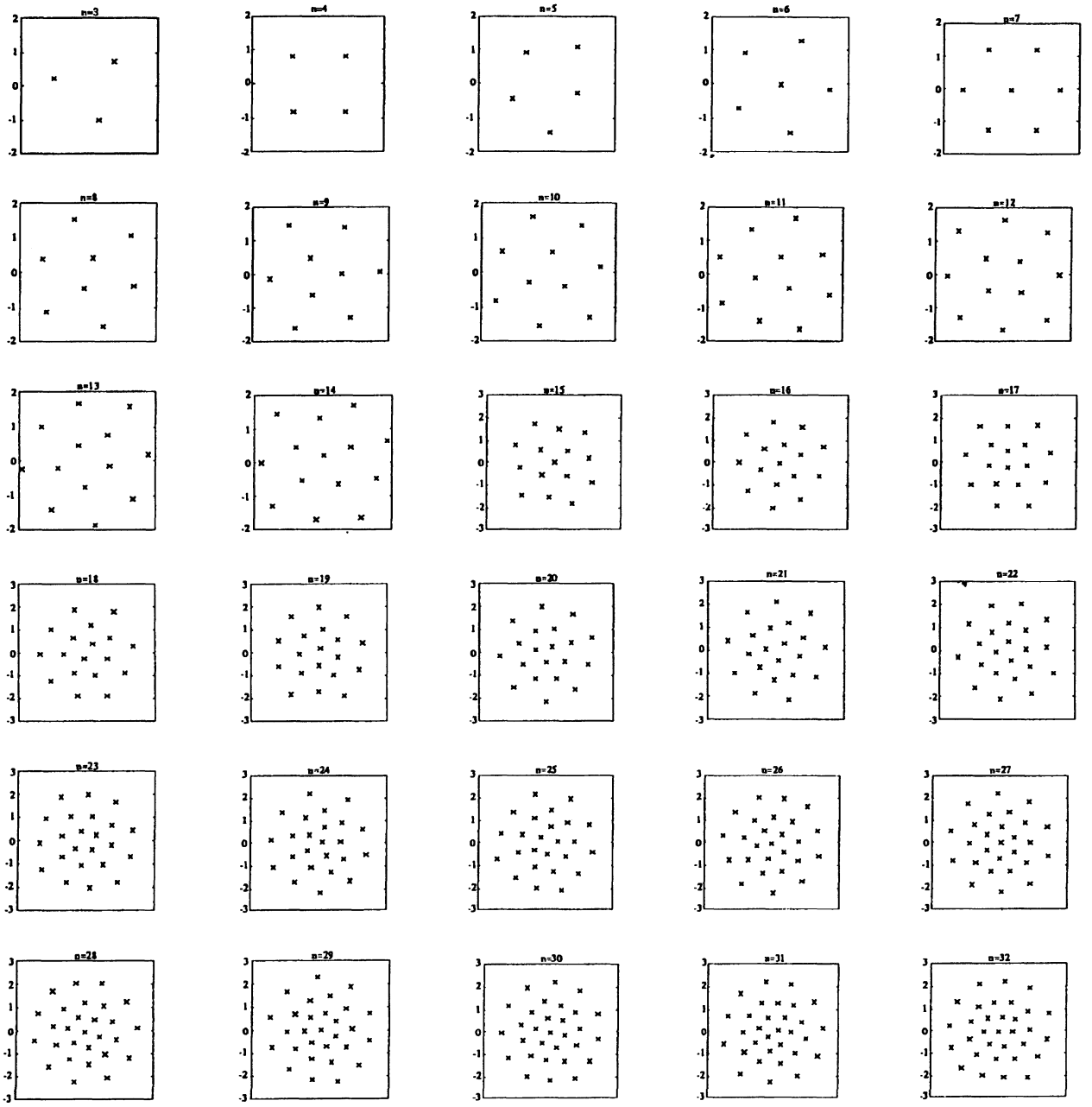

Figure A